REVISION LIST (Topical Editor)

Title:	Extending the Coverage Area of Regional Ionosphere Maps Using a Support Vector Machine Algorithm
Authors:	Mingyu Kim and Jeongrae Kim
Date:	December 18, 2018

Dear topical editor,

Thanks for your comments on this manuscript. The authors have incorporated all the comments in revised manuscript. The revised or new sentences are colored in red in the revised manuscript. The manuscript has been revised from the third revision posted on November 19 (Response to Referee #3), which incorporated all comments from three referees. Previous responses to referees #1, #2, and #3 are attached after this response letter.

< Editor >

1) Page 2, paragraph 25: The TEC variation is correlated with the diurnal and seasonal time variation, and the ionospheric delay above the locations involved in the study reaches its maximum around 14 hours local time (LT) and its minimum around 2 LT. Also, the TEC is higher in spring and autumn, and lower in summer and winter (here you mean the same local time?).
Thanks for the correction. The sentences have been updated. The TEC variation in the second sentence implies a daily mean TEC variation. "daily mean" has been added to the TEC. <Sec.2, p.2>

2) I agree with the referee's comment that using both F10.7 and SSN is not necessary, or you need to argue the necessity/importance of using both indices.
Two additional set of estimations have been performed; (a) with F10.7 only and (b) with SSN only. The authors had already performed these estimations for answering referee #2's comment. This time more comprehensive runs have been performed. Each NN/SVM estimator has been configured as an optimal structure after series of parameter tuning. The revised results still show that the dual parameter case (F10.7+SSN) outperforms the single

cases (F10.7 only or SSN only). The variation of F10.7 and SSN are very similar, but they are not the same quantity. The variation of two parameters are somewhat different and their complimentary aspect may improve the estimation performance. One of the advantages of machine learning is automatic determination of parameter weightings, and a high correlation between input parameters is not a significant problem as conventional linear regressions. Another aspect is the limitation of the data length. The results are based on one-month data. If we process multiple years of long data under different space environment, the results might be changed.

The estimation results of two single parameter cases, with F10.7 only and with SSN only, are added at the end of Section 5. Full description of the results are at the end of this letter.

<Sec.5, pp.14-15>

3) As for using *Kp* and *Dst* indices, there it is necessary to take into account different sources of ionospheric disturbances CMEs and CIR/CH HSSS. CMEs induce non-recurrent storms, while recurrent storms are driven by high-speed solar wind and reappearing with about 27-day periodicity, when the same coronal hole (CH) is facing the Earth. During this kind of disturbance the *Dst* index remains smaller, but because fast streams with southward IMF component may last much longer, the CIR/HSSS related storms have a longer duration, and the cumulative effects of these storms could be more severe than the effects of CME-related storms with significant decrease in *Dst*. (Buresova, et al., 2014). So in the case of coronal holes you need to monitor polar activity (e.g. *Kp, AE* , unfortunately, is not available in real time)

Thanks for your information. The difference between Kp and Dst may explain the estimation results using Dst (instead of Kp), which was presented in the letter to Referee #3's comments. Using Kp was better than Dst in the estimation results. Much longer data span, e.g. multiple years, may be required to analyze the Kp and Dst parameter effects. It can be a good for further research. Sentences are added to discuss this aspect.

<Sec.2, pp.2-3> <Sec.4, pp.10> <Ref., pp.16>

< Supplement for comment #1 >

(a) Parameter optimizations for two single parameter cases (F10.7 or SSN)

Estimation accuracy by NN or SVM can be affected by the design of estimator structures;

parameter value C or number of layers, etc. Optimal estimator structure is changing with the selection of input parameters. Before comparing the single parameter (F10.7 only or SNN only) results with the dual parameter (F10.7 and SNN) results, series of parameter optimizations have been performed for determining an optimal estimator for NN or SVM.

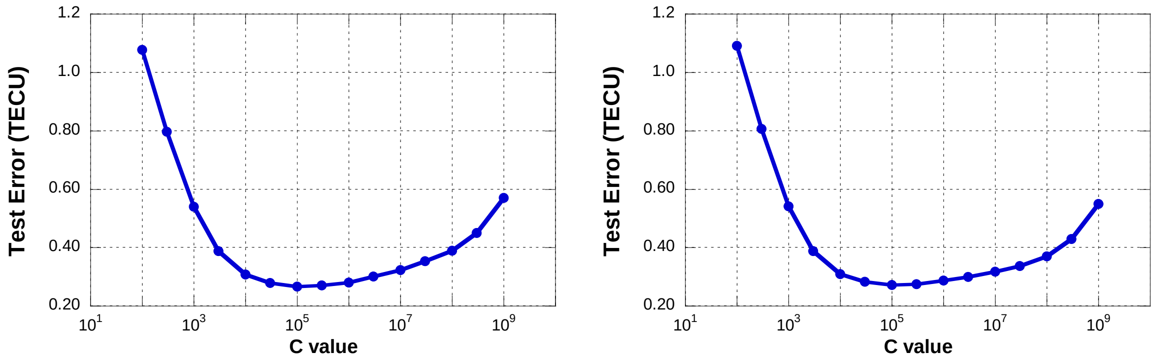

Figure 1. SVM test errors of different C values at S5; F10.7 only (left) and SSN only (right)

Figure 1 shows the optimization results of SVM models with F10.7 or SSN. The optimal C values are computed as the lowest mean RMS errors on the 5˚ extrapolation points. For the both single parameter cases, an optimal C value is determined to 10000.

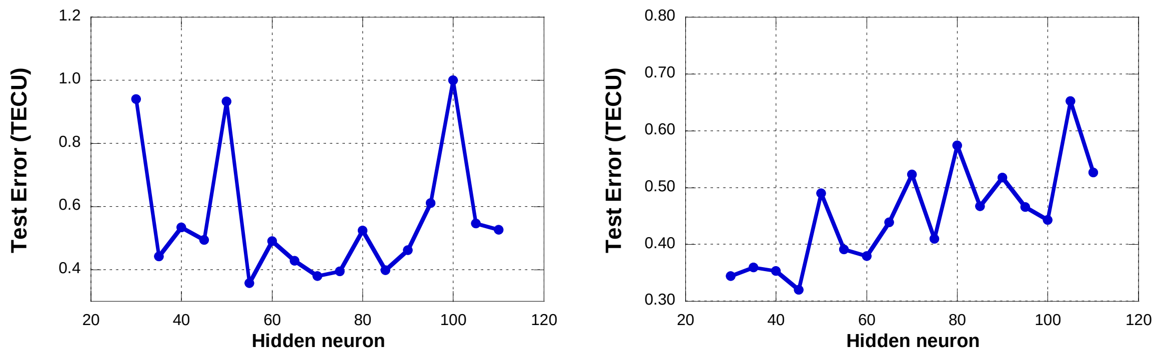

Figure 2. NN test errors of different number of neurons at S5; F10.7 only (left) and SSN only (right)

The optimization results of NN model for each F10.7 and SSN are shown in Fig. 2. The optimal numbers of hidden neurons are 55 for the F10.7 case and 45 for the SSN case.

(b) Comparison among three cases (F10.7, SSN or F10.7+SSN)

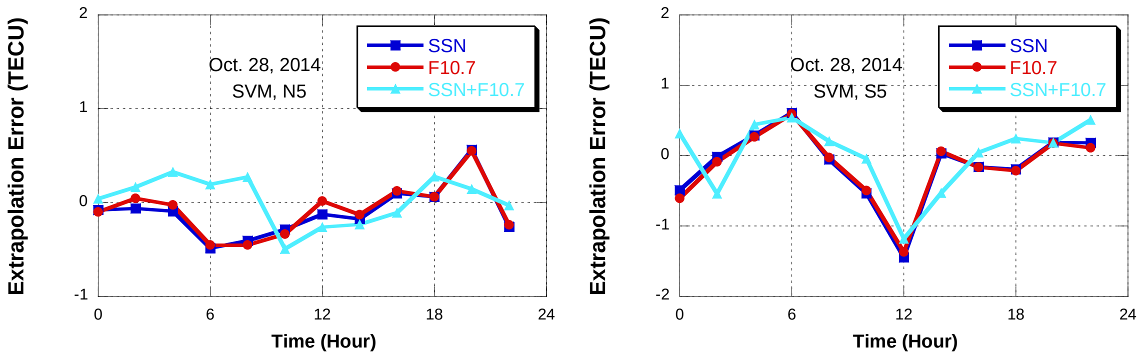

Figure 3. Extrapolation error variations on October 28, 2014 (N5 and S5) using SVM

The estimation results of SSN only and F10.7 only are very similar. The overall error magnitude of SSN+F10.7 at 6-10 UT is smaller than the two single parameter cases.

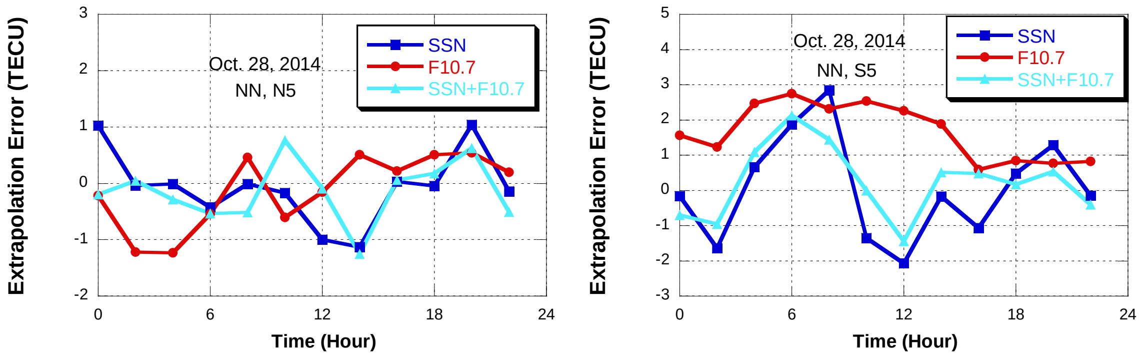

Figure 4. Extrapolation error variations on October 28, 2014 (N5 and S5) using NN

Unlike the SVM results, there is a large difference between F10.7 and SSN when using the NN. The estimation error magnitude of SSN+F10.7 is smaller than that of the single parameter in nighttime (18-02 UT).

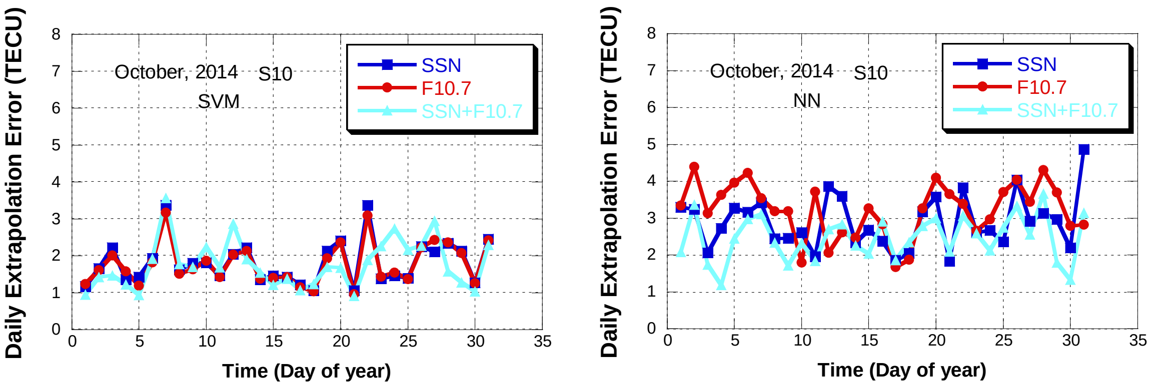

Figure 5. Daily extrapolation RMS error variations in October 2014 (S10) using SVM (left) and NN (right)

The daily estimation RMS errors at south 10˚ point for one-month are compared in Fig. 4. Kp value was significantly increased to 5 at 12:00 UT on October 2 and Dst was decreased to -34 nT on October 1. At this time, the estimation errors of SSN+F10.7 are smaller than that of the single parameter cases. Especially, the errors of SSN+F10.7 are also smaller when CIR-driven geomagnetic storm occurred on October 19-22. During this period, the estimation error reductions with the dual parameters are 26% for SVM, and 22% for NN. It implies that F10.7 and SSN may be complementary to each other during high ionospheric activity period.

Table 1. One−month mean of extrapolation RMS errors of SVM model (unit= TECU)

Extrapolation region	5˚			10˚			15˚		
	F10.7	SSN	Dual	F10.7	SSN	Dual	F10.7	SSN	Dual
North	0.31	0.31	0.32	1.11	1.14	1.02	2.05	1.98	1.97
East	0.26	0.25	0.25	0.58	0.57	0.51	1.06	1.06	1.00
West	0.26	0.26	0.24	0.78	0.80	0.64	1.25	1.26	1.27
South	0.66	0.67	0.67	1.95	1.93	1.89	3.62	3.65	3.58
Total	0.34	0.34	0.33	1.11	1.11	1.01	2.00	1.98	1.95

Table 2. One−month mean of extrapolation RMS errors of NN model (unit= TECU)

Extrapolation region	5˚			10˚			15˚		
	F10.7	SSN	Dual	F10.7	SSN	Dual	F10.7	SSN	Dual
North	0.66	0.64	0.68	1.16	1.20	1.06	2.05	2.01	1.90
East	0.25	0.24	0.20	0.73	0.68	0.71	1.14	1.15	1.13
West	0.25	0.23	0.25	0.78	0.83	0.63	1.46	1.48	1.44
South	0.88	0.73	0.67	2.82	2.63	2.54	3.90	3.91	3.79
Total	0.51	0.46	0.45	1.37	1.34	1.23	2.14	2.14	2.06

The extrapolation RMS errors of the single (F10.7 or SSN) and dual (F10.7+SSN) parameters are presented in Table 1 (SVM) and Table 2 (NN). The total mean errors of the single parameter cases are greater than the dual parameter case at all extrapolation points for both estimation models. Increase of the NN errors with the single parameters at North and South points are significant. Effect of F10.7 and SSN may be complementary to each other during geomagnetic storm days (October 19-22). In this period, the estimation error reduction by the dual parameters are 26% for SVM model and 22% for NN model.

REVISION LIST (Referee #1)

Title:	*Extending the Coverage Area of Regional Ionosphere Maps Using a Support Vector Machine Algorithm*
Authors:	*Mingyu Kim and* Jeongrae Kim
Date:	*October 25, 2018*

Dear Referee #1

Thanks for your comments on this manuscript. The authors have incorporated all the comments in revised manuscript, which are very helpful to improve the manuscript. The revised or new sentences are colored in red in the revised manuscript.

< REVIEWER 1 >

1) *In the introduction (before the reference of Leandro and Santos) you write: "Ionospheric delay observations were used as the input parameters, and the TEC outside the coverage area was predicted". Therefore I would assume that the output is TEC, but looking at figure 4 on the y axis is written ionospheric delay. Therefore it would be better to eliminate this incongruence writing . . ... and the ionospheric delay outside the coverage area was predicted.*

2) *In section 2, Parameter modelling, you write "An extrapolated ionospheric delay, TECext, may be represented. . .." Why you label the ionospheric delay as TECext? Again, after Eq. (3), you talk about ionospheric delays but then in Eq. (4) you refer to observational parameters as TEC obs. It seems that you identified ionospheric delay with TEC. This sounds me strange because, to be meticulous, the ionospheric delay is proportional to the TEC therefore they are not exactly the same thing. To avoid this "incongruence" I suggest to replace in Eq. (4) TEC1 obs with ID1 obs, TEC2 obs with ID2 obs, TECN obs with IDN obs.*

3) *In section 3, subsection 3.1 again you write "y is the target that represent the true TEC in the extrapolation region". For the reasons written above I suggest to replace true TEC with true ionospheric delay.*

TEC values were used as the model parameters and it is why TEC terminology is used in the manuscript. However, all the TEC has been replaced with "ionospheric delay (ID)" to avoid possible confusion. Thanks for your tip.

<Sec.1~6>

4) In section 3, subsection 3.1, in Eq. (5), in the term WT the meaning of T is missing, T stands for? Moreover in Eq. (5) x should be replaced with xSVM to follow the same nomenclature adopted in the flow chart of figure 1 where you write xSVM = [xt xe xobs].

We have added the description of wT where T stands for 'transpose'. Also, we have replaced 'x' with 'xSVM'.

<Eq.5 & 6, p.3>

5) You write, at page 4 after figure 1, "Targets include the true TEC in the j-th extrapolation point", and looking at the flow chart of figure 1, I read ySVM = TECJext.. Therefore I ask you if it is not the case to replace in Eq. (6) and subsequently in the text, y with ySVM .

We have changed "y" to "ySVM".

<Eq.5 & 6, p.3>

6) At the end of subsection 3.1 you write ". . .using the interior point method", please provide a reference here.

A Reference has been added.

<Sec.3.1> <Ref> (Ferris and Munson, 2004).

7) At page 7 you write: ". . .the high ionospheric delay season is more appropriate when evaluating the extrapolation algorithm than the low ionospheric delay season". This sentence is not clear. Could you explain in other words this concept? What do you mean with "more appropriate ".

When the ionospheric delay (ID) values and its variations are small, the extrapolation values become small, and the extrapolation errors become small as well. For the performance evaluation of the extrapolation algorithm, large extrapolation errors are preferred for comparison. Sentences are added for explaining this issue.

<Sec.4, p.7>

8) Looking at the caption of figure 4, I note some incongruences : a) you write "One year variation. . ..(October 01, 2013 to October 30, 2014)" . This period is made by 13 months so it is a year + one month. Probably it would be better to write the caption as: Ionospheric delay for the training period (01 October 2013 - 30 September 2014), and prediction period (01 - 31 October 2014). This is compatible with what you write before figure 4, i.e, "The training period is set to one year from October 1, 2013 to September 30, 2014."and "The prediction period is set to one month from October 1 to 31, 2014".

Thanks for your comment. We have changed the text label.

<Fig.4, p.7>

9) At the beginning of Section 5.1 you write "the variation of the TEC and the. . .". But given that the figure 5 shows on the y axis the ionospheric delay, again, I think it would be better to write: "ionospheric delay variations and the extrapolation results are analyzed for the data from October 28 2014, when. . .. "

We have changed the "TEC" to "ionospheric delay"

<Sec. 5.1, p.9>.

10) In figure 5 the acronym BRDC is not defined, it should be defined in the caption.

BRDC is the acronym of "GPS broadcast message" and represents the GPS Klobuchar model from GPS broadcast ephemeris. We have replaced "BRDC" with "Klobuchar" to avoid possible confusion.

<Fig.5, p.8>

<Table 2, p.12>

11) After looking at figure 5, showing the ionospheric delay trend for the direction north (N5) and south (S5), I would have expected to see in figure 6 the extrapolation error trend still for the direction north and south, but strangely you provide the results for the east direction (E5). Why?

E5 results were selected because its error variation pattern is different from N5. However, for the consistency S5, we have replaced the E5 results with S5 results. Thanks for your suggestion.

<Fig.5, p.9>

12) At pages 8 and 9 you provide the coordinates only for N5 and S5, then you mention the points E5, W5, N15, S15, E15, W15 without giving any info about their coordinates. For completeness, I suggest to delete the info about the coordinates of N5 and S5, providing however a table where are reported all the coordinates of the key points involved in this analysis.

A table of the key points has been added to list the coordinates and related sentences have been revised.

<Table 1, p.7>

<Sec.5.1, p.8>

13) For completeness, in addition to figures 7 and 8, I suggest to insert also a new figure showing the results for 10° extrapolation regions.

A new figure and paragraph presenting the results for 10 degree extrapolation points have been

added. *<Fig. 8, Sec.5.1, p. 10>*

14) *With regard to Table 1, you write: "In the east and west points. . .. . . but the improvement increases as the distance of the extrapolation region increases". If I look at the errors for the direction east and west for SVM and NN, I note the following differences: 0.03, 0.20, and 0.13 (east direction), and 0.01, 0.01, and 0.17 (west direction) for 5◦ ,10◦ , and 15◦ respectively. So, to be meticulous, it is not fully true what you say, because in the east direction 0.13 < 0.20 and in the west direction practically there is not difference between 5◦ and 10◦ . Therefore I suggest to delete the sentence "but the improvement increases as the distance of the extrapolation region increases".*
Thanks for your suggestion. The sentence has been deleted. *<Sec.5.2, p.12>*

(References) *a) The reference Kim and Kim (2014) is in the text, but it is missing in the list of reference* *; The reference year in the list has been corrected from 2014a to 2014.* *<Ref., p.13>* *b) The reference (Kim et al., 2014) is in the text, but it is missing in the list of reference* *c) The reference Kim et al., 2014b is in the list of references, but it is missing in the text* *; The reference year has been corrected from 2014b to 2014, and corresponding sentence has been changed from (Kim et al., 2014b) to (Kim et al., 2014).* *<Sec.5.1, p.8> <Ref., p.13>* *d) Please replace in the list of references McKinella with McKinell* *; Corrected. Thank you. <Ref., p.13>*

REVISION LIST (Referee #2)

Title:	*Extending the Coverage Area of Regional Ionosphere Maps Using a Support Vector Machine Algorithm*
Authors:	*Mingyu Kim and Jeongrae Kim*
Date:	*November 12, 2018*

Dear Referee #2

Thanks for your comments on this manuscript. The authors have incorporated all the comments in revised manuscript. The revised or new sentences are colored in red in the revised manuscript.

< REVIEWER 2 >

1) *The performance of a NN model largely depends on the number of hidden layer neurons used. The authors indicate that they have used 80 hidden layer neurons based on previous studies. The previous study referenced does not give a convincing method to check overtraining of the networks. Also the dataset is entirely different, and the NN architecture is also different. Using a different number of hidden layer neurons may give better results, perhaps better than the SVM method. I therefore suggest that the authors device a system to check performance of the networks (especially on extrapolation datasets), if not, the networks may even over-fit the training data and so perform poorly on extrapolation data. The authors may also choose to indicate/explain in the manuscript that the observation they report is not generalized (but limited to the case of their NN training) because a carefully done NN may give better results, even than the SVM does.*

The parameter optimization of the SVM and NN models had been already performed but not included in the manuscript. We added a description for the optimization results. SVM has some advantages over NN as described in Section 1. As the reviewer pointed out, the advantage of SVM might be reduced in other ionospheric environments, but ultimate generalized methodology is not within the scope of our research. The authors concluded that the SVM outperforms NN in case of our research after series of computation runs.

< Figs. 5 and 6, p.9>

2) *There is also information which appears missing in the manuscript. Inputs for the models do not include station locations? How do the models predict different values for different locations? The spatial structure (with station locations) is pre-fixed in the models? How do you query the models for data of, let's say, 10 degrees from the center of your circle? I wonder what applications there are for this method if the spatial structure for the models is pre-fixed.*
The proposed algorithm is using fixed locations both for input and output, and it does not require the spatial structure. Other researchers' works on ionosphere prediction used raw GPS TEC measurements at varying IPP (Ionospheric Piercing Point) and the measurement locations should be registered in the input. Our algorithm uses a grid-based ionosphere map with fixed grid points, and their location information is not required as the model inputs. *Our algorithm is to estimate TEC at fixed pre-specified locations. The TEC values at the estimation points are used during training process and their actual location information is not necessary. In order to change the output point, a new training process should be performed. In practice, a gridded ionosphere map is generated from the extrapolation algorithm and the ionospheric delay at specific point, e.g. IPP, can be computed through interpolation. Sentences have been added to emphasize this aspect.* *<Sec.2, p.3>*

3) *Although the authors have used data for South Korea, they do not indicate the implication of this limitation anywhere on the manuscript. Given the spatial variability of the ionosphere, extrapolation schemes for a given region will perform differently for different regional models. For instance, whether the ionospheric ionization should be greater or otherwise in the outer regions is something too arbitrary to decide based on the inner data. And if the outer data will always be required to train the relationship, then the application I see of this work is defeated.*
As the reviewer pointed out the ionosphere environment depends on its geomagnetic locations, and our extrapolation algorithm performance might be different at other locations. However, our proposed algorithm is still worth for extending ionosphere map coverage. If the estimation region is changed, a new training and optimization process should be performed. Arbitrary variation of outer points may limit the usability of this algorithm, but it does not imply that this algorithm is useless. In any case, ionosphere variation has some sort of geographical correlation and it is the key concept of this algorithm. Paragraphs on the limitation on the algorithm have been added. *<Sec.5, p.14>*

4) *Page 2, lines 32-33: It is not clear why two solar activity indicators (F10.7 and SSN) are repeated. Also, how does the method in this work take care of the time lag (up to several hours/days) for geomagnetic storm effects to be observed in the ionosphere?*

The extrapolation accuracy with both F10.7 and SSN was better than with F10.7 only in some cases. At S5 region, the 4.4% and 2.9% error reduction was achieved for SVM and NN model, respectively. Based on this experiments, we included SSN in environmental inputs.
For geomagnetic storm, Kp value is included in the input. Although even Kp may have several hours of delay, the ionosphere estimate is much more dependent of its inside ionosphere map input, which responds to geomagnetic storm in real time (since the inner ionosphere map is the output of real-time ionosphere measurements).

<Sec.2, p.2-3>

5) *Page 1, lines 37-38: "Kim and Kim (2016) additionally used ionospheric delays in the inner ionospheric coverage area." It is not clear what this sentence means, and why it is necessary to include it here.*

This sentence was used to emphasize spatial extrapolation rather than time series prediction. The sentence has been revised to clarify.

<Sec.1, pp.1-2>

6*) Page 3, line 16: "In the above equation..." should read " In equation 7...".*

7) *Consider using "ionospheric map/model" in places of "ionosphere map/model" throughout the manuscript.*

We have changed all the corresponding words. Thanks for the tip.

<Sec.3.1, p.3>, <Sec.1-6>

8) *Page 7, line 1: Authors should clarify what previous one-epoch values are referred. What is the interval between successive epochs? Is the interval between successive epochs sufficiently small for previous one-epochs to be safely used? And what happens if there may be no data for previous one, two, three. . .. epochs?*

The estimation interval should be the same as the inner ionosphere map input interval. In this research, two-hour interval was used because two-hour interval IGS global map is implemented for the inner map. If a shorter interval inner map is used, e.g. 5 min. SBAS map or real-time GPS-derived map, and then the estimation interval becomes shorter. Our algorithm is not a time-prediction algorithm, as other preceding researches, and the estimation interval is not an important factor to determine the accuracy. It is also because that the estimate more depends on inner ionosphere map than other environmental parameters, e.g. F10.7, Kp etc.

If the inner ionosphere map is complete unavailable, our algorithm should not be used because it is not a prediction algorithm but an extrapolation algorithm. If some points in the inner ionosphere map are not available due to data loss or availability, interpolated data from near grid points replace the data at those points.
If some of environmental parameters, e.g. F10.7 and SSN, are unavailable or delayed, time-predicted value of those parameters can be used. We have already tested prediction of the environmental parameters, but the effect of prediction was not significant because the inner ionosphere map takes a dominant role in the estimation. Sentences have been revised to clarify this aspect.

<Sec. 4, pp.7-8>

9) *The authors cite SVM applications to other fields but not a citation on previous ionospheric applications. There have been previous studies on the use of SVM for Ionospheric research. E.g.:*
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2010RS004393
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2010RS004633
https://www.ann-geophys.net/31/173/2013/angeo-31-173-2013.pdf

Thank for your tip. Overview of the three papers has been added.

<Sec.1, p.1>, <Ref., pp.14-15>

REVISION LIST (Referee #3)

Title:	*Extending the Coverage Area of Regional Ionosphere Maps Using a Support Vector Machine Algorithm*
Authors:	*Mingyu Kim and Jeongrae Kim*
Date:	*November 19, 2018*

Dear Referee #3

Thanks for your comments on this manuscript. The authors have incorporated all the comments in revised manuscript, which have been helpful to improve the manuscript. The revised or new sentences are colored in red in the revised manuscript.

< REVIEWER 3 >

***1)** In page 7 line 2 you say that correlation coefficient between the two adjacent data for F10.7 cm, Kp and SSN is 0.93, 0.863, 0.852. I think that correlation analysis should be between F10.7 cm-TEC, Kp-TEC, SSN-TEC. According to these results, you will do weighting and this weighthing affects your results. The other thing is that F10.7 cm is only one value for one day, therefore its affect can not be monitored effectively in a day. You can obtain Kp value from https://omniweb.gsfc.nasa.gov/form/dx1.html with one hour resolution. You should also take into account DsT index which give information about geomagnetic activity of ionosphere. You can also more precise results by adding this DsT index.*

(a) The correlation analysis of TEC with DsT, F10.7, Kp, or SSN, has been recently performed, and details results are presented in the supplement of this revision list. Due to the complicated nature of the ionosphere correlation, the correlation is not clearly shown in the plots. However, Dst shows a more correlation than the other parameters as the reviewer suggested. Since the NN or the SVM uses a complicated weighting structure (rather than a linear weighted combination), individual weights cannot be assigned in the algorithm. Sentences on this correlation analysis have been added.
<Sec.2, p.2> <Sec.4, pp.9-10>

(b) The 1-hour Kp index that the reviewer suggested had been investigated. At this moment, they provides only 3-hr data online; 1-hour Kp data seems to be the same as 3-hr data, i.e. simple extension. Our estimation algorithm's dependency on the environmental parameters, e.g. Kp or F10.7, is much less than the inner ionosphere data. The effect of Kp update rate may be more

important during geomagnetic storm, and it can be investigated as a further research topic. Discussion on this environmental parameter update rate had been added in the second revision of the manuscript.
<Sec.4, p.8, lines 1-12>

(c) Based on this correlation analysis, we have rerun all the estimation process after replacing Kp with Dst. The detailed estimation results are presented in the supplement. Our preliminary results show that Dst is not better than Kp for the estimation. These results may be due to incomplete optimization for Dst use or Dst nature. Optimization with Dst can be good as a further research topic. The Dst results (Figures and table) have not been added to the revised manuscript, but a discussion on these experiments has been added.
<Sec.4, pp.9-10>

2) In this paper, I cannot see any discussion, therefore your results can not be confirmed. Please investigate other studies and compare your result.

We have added more discussion for analyzing the results after reviewing additional papers on machine learning ionosphere researches. However, no one have performed similar concept of research (the ionospheric delay extrapolation using inner data points) as far as the authors' know, direct comparison or discussion with preceding researches was not possible. In depth analysis clarified the correlation between the ionosphere spatial gradient and the estimation accuracy. Sentences on the new discussion and several references have been added.

<Sec.5.1, pp.11-13> <Sec.5.2, p.14> <Ref., pp.15-16>

3) Conclusion also should be explained detaily. You mention about differences but you did not any comment on this. Please explain these differences.

We have added some paragraphs on the ionospheric delay gradient which explains the high estimation error at the north point unlike other ionosphere estimation researches. For explaining the accuracy difference between the SVM and the NN, the inherent advantage of the SVM over the NN.

<Sec.6, p.15>

< Supplement for comment #1>

(a) Correlation analysis

Correlation plots between TEC and environmental parameter are generated using one year of data (from Oct. 1, 2013 to Sep. 30, 2014). The TEC value at S5 point is used for the analysis, and four parameters, Dst, Kp, F10.7, and SSN, are considered.

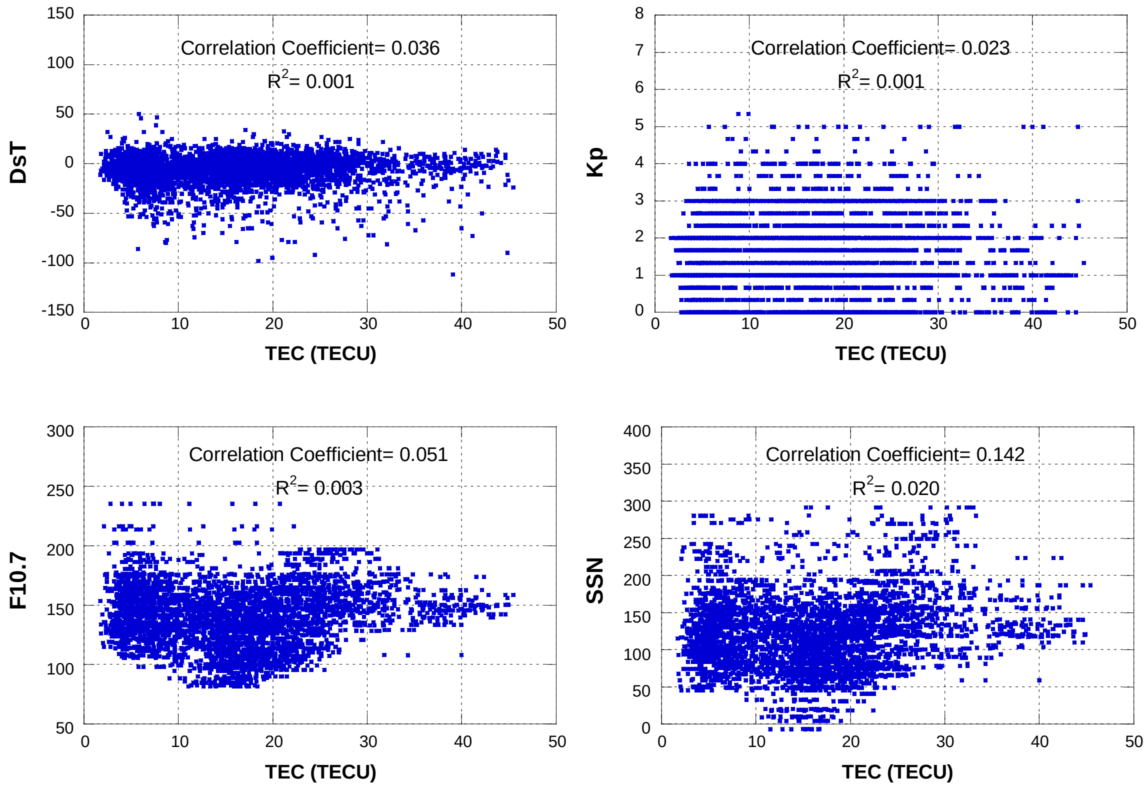

Figure 6. The correlation between ionospheric delay and Dst, Kp, F10.7, and SSN

The correlation coefficients of Dst, Kp, F10.7, and SSN are 0.036, 0.023, 0.051, and 0.142, respectively. The correlation coefficient is lower than expected, but Dst shows a slightly larger value than Kp as the reviewer suggested. Based on this correlation analysis, another set of experiments had been performed with Dst. Since the TEC estimate is obtained from complicated non-linear weighted sum of the environmental parameters, a linear correlation does not directly affects the estimation accuracy.

(c) Experiments with Dst

Test set of experiments had been performed by using Dst instead of Kp. All the figures and tables were re-generated using Dst, but few figures are selected for the comparison. Some plots have different y-axis limits.

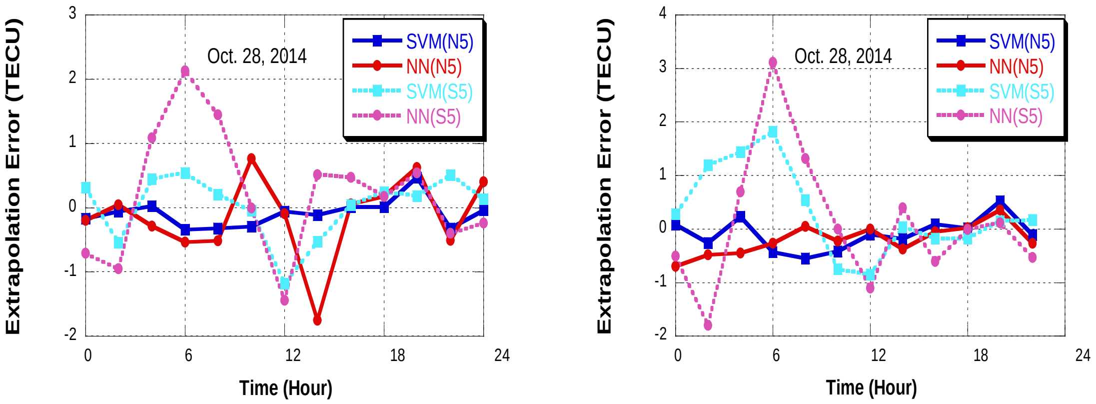

Figure 7. Extrapolation error variations on October 28, 2014 (N5 and S5) using Kp (left) or Dst (right)

In overall, the estimation errors are increased with Dst by both of the SVM and the NN. Especially the error increase at peak ionosphere at 6 UT is significant

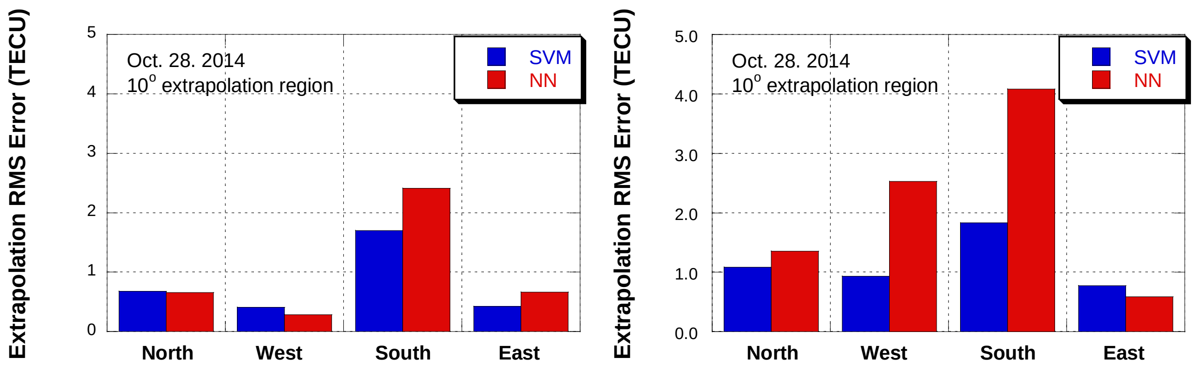

Figure 8. Extrapolation error for each direction (10° point) using Kp (left) or Dst (right)

The estimation error at 10° points are compared in Fig. 3. The overall errors (both of the NN and the SVM) are increased with Dst. The error increase by the NN is more significant than the SVM. It also proves that the NN is less robust than the SVM or more sensitive on input parameters.

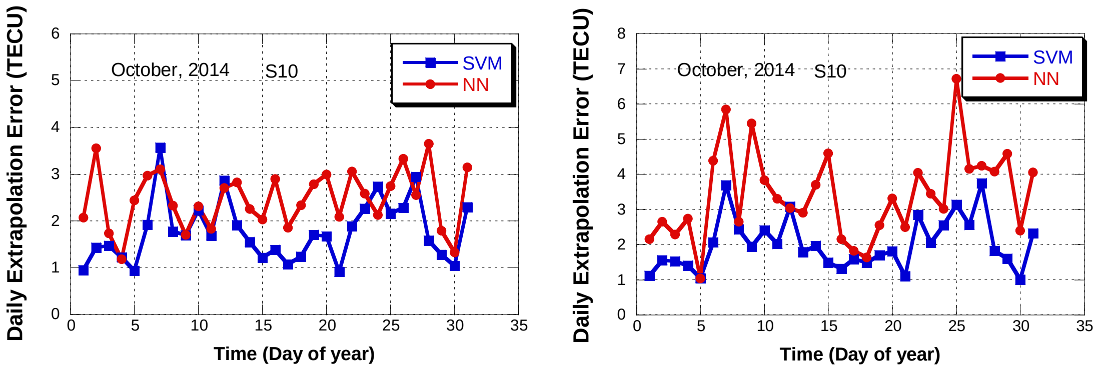

Figure 9. Daily extrapolation RMS error variations in October 2014 (south 10° point) using Kp (left) or Dst (right)

The daily estimation error variations during one-month are compared in Fig.4. The errors of a SVM model is slightly increased, but the errors of a NN models is significantly increased when Dst < -25nT (October 9, 19-21, 28). It means that when the geomagnetic activity is high, the extrapolation performances of a NN model with Dst might be poor unlike the SVM model. Although the geomagnetic activity is low in October 9, the error of a NN model is also increased about 2.70 TECU.

Table 1a. One−month mean of extrapolation errors with Kp (unit= TECU)

Extrapolation region	5°			10°			15°		
	Klob.	SVM	NN	Klob.	SVM	NN	Klob.	SVM	NN
North	14.41	0.32	0.68	13.07	1.02	1.06	12.04	1.97	1.90
East	14.63	0.17	0.20	14.57	0.51	0.71	14.47	1.00	1.13
West	13.38	0.24	0.25	13.29	0.64	0.63	13.12	1.27	1.44
South	25.13	0.57	0.67	24.40	1.89	2.54	26.97	3.58	3.79
Total	16.89	0.33	0.45	16.33	1.01	1.23	16.65	1.95	2.06

Table 2b. ***One-month mean of extrapolation errors with Dst (unit= TECU)***

Extrapolation region	5°			10°			15°		
	Klob.	SVM	NN	Klob.	SVM	NN	Klob.	SVM	NN
North	14.41	0.37	0.60	13.07	1.32	1.59	12.04	2.40	2.45
East	14.63	0.37	0.36	14.57	0.83	0.87	14.47	1.28	1.51
West	13.38	0.34	0.36	13.29	0.79	1.15	13.12	1.47	2.97
South	25.13	0.69	1.21	24.40	2.13	**3.62**	26.97	3.78	3.93
Total	16.89	0.44	0.63	16.33	1.27	1.81	16.65	2.31	2.72

The extrapolation RMS errors for all directions and models are compared in Table 1a and 1b. The values changed by using Dst instead of Kp are shown in red. All the errors are increased except of NN error at N5 point. In S10 with NN model, the error is significantly increased because of the poor estimation in period of Dst < -25 nT.