Reviewer comments:  
  
Reviewer: 1  
Comments to the Author  
Thank you to the authors for writing a clear and well-organized paper, especially considering that their native language is not English.  
  
This manuscript addresses two issues:  
  
1) describes a new method to directly measure [DON] in freshwaters, where [DON] are relatively high (i.e., > 0.5 mg N/L, or 35 µM)  
  
2) quantifies types of error associated w/ calculation of [DON] from [TDN] measurements, after subtraction of [DIN]  
  
The first part of this paper is very strong; the authors have comprehensively tested their new method and shown that in freshwater systems with high [DON], they can reliably separate DIN from DON. This method deserves publication and will likely make a significant contribution to studies of terrestrial systems, especially anthropologically-influenced systems with high [DON].  
  
The second focus of the manuscript, identifying causes of error associated with traditional (i.e., HTCO) [TDN] measurements, is a noble goal, but the results are ambiguous (see top of page 13), and it is not clear to me that the authors have demonstrated anything not already known (indeed, they summarize the state of this analysis in the introduction and not much seems different there from what is presented in the results). As they summarize at the top of page 16 and in Figure 3, while the coefficient of variation (a percentage) does not increase with increasing DIN or TDN of a sample, the magnitude of the error in the calculated [DON] will increase as that constant fractional error associated w/ [TDN] analysis is multiplied by an ever increasing [TDN] (as a function of increasing [DIN]). To me, this seems to be what we have known all along, but does not get at a mechanistic understanding of what goes wrong in [TDN] sample analysis when DIN is >50% of a TDN sample.  
  
Moreover, given that the model compounds used in this study to do not behave the way natural organic matter behaves (see Figure 2a), this error analysis seems to be further compromised; it is not relevant to natural samples. Similarly, on p. 24, lines 38-43, the authors state that: “Although these errors were systematic, they occurred in different directions and amounts for the different samples tested in this study.” – to me, this indicates that in fact these errors are not systematic. How could one predict or systematically correct for them if they have different magnitudes and signs in different samples, and the model compounds don’t behave as natural samples do? Please remove the text, “although these errors were systematic”, or else describe in more detail how these errors are systematic.  
  
I suggest that the authors focus on the methods aspect of this manuscript, and reduce text (and perhaps figures, i.e., Fig 2b, Fig 3) associated with these statistical tests – they end up distracting from the new method that they have systematically tested and that is deserving of publication.  
  
Major concerns:  
Does this method work for high molecular weight compounds? What is the “size exclusion”? Is there a molecular weight cutoff? Please state explicitly in the methods and discussion what range of molecular weights this method applies to. I realize the authors reference other methods papers, but stating this explicitly here will make this manuscript more accessible and useful to readers.  
  
Regarding the model compounds used, I found them to be not significantly different from each other, and not particularly challenging examples of DON to oxidize. Tyrosine should be relatively easy to oxidize – there is an ammonia group hanging off of a carbon chain. Similarly, imidazole does not appear to be a particularly demanding compound to oxidize – it is a pentacycle with two nitrogen atoms. Same for nicotinic acid and glycine I would have liked to see one a compound like antipyrine included in the tests (see Bronk et al., 2000).  
  
Figure 2a: Shouldn’t it be a concern that the model compounds do not behave as the natural organic matter samples behave? Shouldn’t that indicate that either the model compounds are not good “models”, in that they don’t behave like natural samples, and/or that there are other components of natural samples (i.e., DOC?) that influence sample analysis? On p. 14, lines 8-13, how do the authors know that DIN concentrations were underestimated in natural samples, and that that caused overestimation of [DON] in natural samples, and conversely, that DIN was overestimated in model samles, leading to [DON] underestimation? Why did [DIN] analysis/measurement bias vary depending on sample type? I think this deserves additional attention in the text.  
  
Minor concerns:  
p. 7, lines 27-32: It is not clear to me that after anion exchange or dialysis that DON determination would be “indirect” – if all the NO3, NO2, and NH4 are removed, then isn’t that a direct analysis of DON? If the authors are referring to samples with high NH4 concentrations, it might help to specify so.  
  
p. 10, lines 54-58: I didn’t understand this sentence.  
  
p. 18, line 52, I suggest inserting the word “relatively” before “low DON concentrations”; again, 0.5 mg N/L = 35 µM DON, which is a concentration that exceeds any found in the ocean by a factor of ~5.  
  
p. 24, lines 51-53: Similarly, I would suggest adding “relatively high DON concentrations and” after “measurement in samples with”  
  
p. 25, lines 46-53: Again, typical [DON] in marine surface waters are ~5 µM (< 0.1 mg N/L), and decrease with depth; this is below the stated detection limit of this method (0.1 mg N/L). Moreover, deep [DON] are between 2-3 µM (we think, although there is significant uncertainty associated w/ these numbers as described in this manuscript); it is not clear that this method has the precision necessary to differentiate between 2.0 or 3.0 µM DON in the deep ocean. I would remove the sentence that this method should in principle be applicable to marine samples until this method has been shown to have the sensitivity required of marine samples and their comparatively low [DON].

Reviewer: 3  
Comments to the Author  
This paper presents a decent argument based on statistical grounds that the current methodology for the measurement of dissolved organic nitrogen (DON) in certain aquatic environments (i.e. where the dissolved inorganic component of nitrogen exceeds the dissolved organic by a factor of 1.5 or more) is seriously lacking.  It proposes the use of size exclusion chromatography (SEC) to separate DON from DIN to significantly improve the detection limit and the random error associated with DON  measurements in such environments.  It performs comparisons of the currently accepted HTCO-TDN difference method with an SEC-DOC method that clearly illustrate the benefits of SEC and its potential to improve DON measurements.  
  
Both the limnological and oceanographic communities have recognized the importance of DON in global nitrogen cycling for more than two decades.  But the community's inability to make measurements of DON to the same degree of precision as dissolved inorganic nitrogen (DIN) has seriously limited interpretation of DON data-sets in many environments where DIN concentrations are larger than DON.  Consequently, the scope of DON measurements has been primarily limited to those aquatic environments where DIN concentration is low to negligible.  This leaves a very large proportion of aquatic environments (both fresh and marine) relatively unexplored in terms of their DON cycling.  
  
By quantitatively addressing the limitations of current DON measurement technology and by testing an alternative method that can seriously improve DON measurement precision for freshwater work (and potentially marine work), this paper deserves publication in L&O:Methods.  However, the paper requires some major revision to help clarify the various measurement "treatments"  that were performed and the interpretations of the measurement  and simulation data.  Overall, I do not feel these revisions change the major argument of the paper, but I feel they are necessary to improve the paper and to better convey the author's message to the reader.  
  
Below I list the requested revisions.  I would be happy to review a revised copy if necessary but am  comfortable deferring to the Associate Editor's judgment that the requests have been satisfied or sufficiently rebutted.  
  
Page 4, Paragraph 2: It would be worth mentioning that inefficiency in converting nitrate to nitrite in the traditional Cd-reduction technique used for DIN measurements can also result in errors in DIN analysis of a few percent or more (if nitrite is present in the samples of course).  These errors have been documented in other studies that measured DON with various indirect, difference methods  -- see Sharp, Marine Chemistry 78 (2002) 171– 184.  
  
Page 5, Line 3: The use of the term "model" substances, while commonplace among many method papers, is misleading.  It implies that the chosen substances (i.e. tyrosine, imidazole, nicotinic acid) mimic the properties of the natural DON in the environment under study.  I don't think the authors intend to claim this.  A more suitable term should be chosen and used throughout to avoid this misinterpretation.  
  
Line 20: The wording "As for random errors, even small systematic errors of the TDN or DIN..." is confusing as it seems to conflate random errors and systematic errors.  This statement should be clarified.  
  
Page 5, Line 50-51: The statement that there is no direct method for DON quantification should be qualified.  Benner, McCarthy and others have utilized ultrafiltration to isolate a fraction (perhaps 20%) of the DON from DIN and then quantified it using catalytic combustion techniques.  This statement should be amended to reflect that fact or the authors should specify they are referring to the bulk pool of DON.  
  
Page 6, Line 20: Specify that the "part" of the DON that would be lost during dialysis would be a low molecular weight fraction.  It is important to introduce the term here as it also could be an issue that affects SEC as a means for measuring the bulk DON concentration (see later comment).  
  
Page 7, Line 33:  Was the synthetic air itself cleaned of potential impurities prior to sparging the samples?  Sometimes small concentrations of ammonia impurities show up in industrial gases (even "high purity air") which can accumulate as NH4+ in the acidified samples during sparging.  
  
Page 7, Line 56: How were the determination limits for the LC-OND determined?  In-house or based on manufacturer specifications.  
  
Page 8, Line 3: The wording "..and the organic part of the dissolved nitrogen is therefore separated from the inorganic part" is confusing.  It implies  that the DON under analysis is a non-charged species which undergoes dissociation to a cation and anion in the chromatographic column.  I don't think that is what the authors intend so this statement should be clarified.  
  
Page 8, Paragraph 1: It would enhance the presentation and interpretation of the SEC results which appear later on in the paper, if a little more information were provided to the reader about the separation of compounds within the column.  Does it separate based on a molecular weight to charge ratio?  How does an uncharged species -- such as a protonated organic acid -- elute from the column?  What is the impact of pH on the sample?  Are these samples acidified (I would guess not but it is unclear).  While I am sure these issues are discussed in the reference provided (Huber et al., 2011b), I would have preferred to see a few more sentences worth of detail on how the column works without having to go to the reference immediately.  
  
Page 8, Paragraph 2: This paragraph would read better if the various equations to calculate DON were each written on a separate line.  
  
Page 9, Line 56-57: The wording "To assess the reliability of the standard approach in dependence on the DIN:TDN ratio of natural samples" is incorrect.  Perhaps the authors meant "To assess the reliability of the standard approach in determining DON over a range of DIN:TDN ratios in natural samples..."  
  
Page 9-12, Samples, Treatment and Analyses section The overall description of the various measurement "experiments" is satisfactory throughout this section (with the exception of one specific comment which follows).  However, I found it rather difficult to keep track of the different treatments throughout the paper and found myself frequently flipping back to this page struggling to search through the text for specific treatment details.  The manuscript would benefit significantly from the addition of a table that could encapsulate the various treatments and act as a ready reference tool for the readers as they go through the results.  
  
Page 11, Line 22-23:  The statement "...recovery rates were calculated as the difference between true concentration of the model substances and the SEC-DON measurements or the standard approach, respectively" is not correct.  This describes the measurement "bias" not "recovery rate".  
  
Page 12, Page 46-47: The conclusion of the sensitivity analysis that the random error in HTCO-TDN dominates the random error of the DON calculation is intuitive.  But the suggestion that there is "little correlation" between variability in DIN and variability in DON for the screening seems like a statistical artifact of the sensitivity analysis.  It seems that a comparison of the variances of TDN and DIN with the variance of DON does not support this for the data presented in the text  For example, assuming the average DIN:TDN=0.75 and DON=1 mg/L, the TDN would be 4 mg/L; the DIN, 3 mg/L.  If all of that DIN is nitrate+nitrite, Variance of DIN = 0.025, Variance of TDN = 0.045, Variance of DON = 0.070.  Does this not suggest that the variability of DIN is responsible for nearly 1/3 of the variability of DON?  How does the screening that was conducted differ from this simple treatment of the data?   It may improve reader confidence in the conclusion if this difference is explained more thoroughly.  
  
Page 13-15 & Figure 2:  For the model substances, it is stated that on average DIN recovery was 103.4% for imidazole, whereas HTCO-TDN was recovered at 99.9%.   This does not seem to match the data presented in Fig 2a.  For example at a DIN:TDN ratio of 0.7 and assuming TDN=2 mg/L (as stated in the treatment description on page 10), the DIN would be 1.4mg/l and DON would be 0.6mg/L.  If DIN is overestimated by ~3.4% then DON ~ 0.55 which is only 90% of the true concentration as opposed to the 70% of true which is plotted in the figure for imidazole.  The same applies to some of the data points plotted for tyrosine.  In order to get the very low DON recoveries for the model compounds that are plotted in this figure, wouldn't the overestimation of the DIN need to be much higher?  Is it possible that a dilution was not being taken into account here? Or was there a range of overestimations of nitrate-nitrite that is not presented in the text?  
  
Page 16, Last few sentences:  I don't think the authors' propose a satisfactory justification for why the sensitivity analysis yields a different conclusion than the relative error analysis of the model compounds.  I am left wondering if the sensitivity analysis is useful or correct since it essentially disagrees with the variance comparison mentioned earlier.  I may be misinterpreting something here, but in the least the authors need to explain this discrepancy more clearly as I am sure other reader's will also be confused and skeptical.  
  
Page 18, Line 6:  The conclusion that systematic errors in DON can be as much due to errors in DIN determination as they are to errors in HTCO-TDN determination has already been noted in inter-laboratory comparisons of DON (see Sharp reference mentioned earlier).  This should be noted somewhere in the text.  A comparison with the results of that paper might also be warranted.  
  
Page 18, Line 28-29: I disagree with the statement that the 0.5mg/L concentration is "close" to the determination limit of 0.1 mg/L.  Five times the concentration of the determination limit is not at all close and the method should easily be able to distinguish such a concentration.  This is especially so if the determination limit quoted for the device is indeed 3x the standard deviation of a signal generated by a measurement blank.  
  
Page 18, Line 43-44: The sentence containing "...the direct HTCO-TDN measurement of DON..." is confusing.  I think the authors mean "...the direct HTCO-TDN measurement for the model compounds..."  
  
Page 18, Line 50-51: Change "...much better than for HTCO-TDN..." to "..higher than the recovery for HTCO-TDN..."  
  
Page 19, Line 36-37: "low-molecular model substances" should read "low molecular weight compounds"  
  
Page 20, Table 2:  It was intriguing that imidazole and tyrosine had similar recovery means for the HTCO-TDN, whereas they differened by nearly 6 percent for the SEC DON.  Was this a significant difference?  Is there a potential for interference from nitrate (MW=62) with low molecular weight compounds like imidazole (MW = 68)?  This could explain why imidazole recovery is higher than tyrosine for any given DIN:TDN ratio and also why there is a general trend toward higher recovery rates for imidazole as the DIN:TDN ratio increases.  That same trend is not present with tyrosine (MW=181).  In my opinion, this potential for interference seems to be the only major potential drawback of SEC-DON.  It would improve the manuscript if the authors addressed this in detail by better explaining the SEC separation, showing examples of model compound retention times for various DIN:TDN ratios and exploring the possibility of interference in natural samples.  I feel this is important since low molecular weight dissolved organic nitrogen compounds such as urea (MW=60) are in measurable and sometimes high quantities in natural waters.  
  
Page 20-22 and Figure 5: It seems very odd that HTCO-TDN would overestimate natural organic matter when the DIN:TDN ratio is greater than 0.4.  Can the authors provide any methodological reason for this considering their specific combustion system?  Is the HTCO system in use somehow more  sensitive to NOx signals generated by oxidized No3 or NH4+.  Were calibrations curves for the HTCO system performed at a range of DIN:TDN ratios?  Were they similar?  Have other authors seen this behavior in their systems.  
  
Page 20-22 and Figure 5:  The comparison between HTCO-TDN and SEC-DON measurements of the natural organic matter seem to be biased in favor of the SEC-DON method.  The NOM that was analyzed was obtained by reverse osmosis so it likely represents a high molecular weight fraction of the dissolved organic matter.  As alluded to in an earlier comment, this would essentially remove the possibility of nitrate interference of nitrate with low molecular weight substances that would elute later in the chromatogram.  Conversely, the HTCO-TDN does not get any advantage since combustion of High molecular weight likely proceeds in a similar fashion as the low molecular weight.  
  
Page 25, Final Paragraph:  I was happy to see the last paragraph on the applicability of the measurement in seawater as I am sure a number of the readers of this article will be oceanographers who are keen to exploit an improved method for DON measurement in the relatively high DIN environment of the ocean. However, I think the authors should add a few more sentences regarding the method's feasibility in marine environments.  How might the presence of salts affect the functioning of  chromatographic column, the inline combustion oxidation system?  How might the presence of larger molecular weight ions (SO4 or Br for example) interfere?