We thank the two referees for their comments and constructive feedback, which we have considered in the revised paper. Their comments (in italics) and our responses are addressed below. All page and line numbers in the referee's comments refer to the originally submitted version of the paper. All page and line numbers in our response refer to the revised paper.

Anonymous Referee #1

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-75, 2018.
Received and published: 26 March 2018

The paper compares two simple methods for estimating error variances from double or triple collocated data using simulated and real data. The error estimation methods are the two and three cornered hat methods (2CH and 3CH). The paper is interesting, but should be considerably improved. The present version is not suitable for publication.

We have revised the paper according to the referee's suggestions.

1. *The presentation is rather sloppy, and it looks more like a report than a scientific paper. For instance, equations (18b-c-d) are essentially the same equation, and indicating cancelling terms by crossing them out (pages 11 and 12) is not the style of a scientific paper. Variance is defined, but MS (equation (2) and further) not.*

We have fundamentally revised the structure of the paper, eliminated some of the equations and intermediate steps of the derivations, moved the 2CH derivation into an appendix, addressed the style issue, and made sure all the terms are defined. We have also eliminated the need for both lower case and upper case X, Y and Z.

2. *More seriously, the authors seem to assume that their data are well calibrated and do not contain any representativeness errors - at least these problems are nowhere mentioned. Therefore they are surprised when pairwise 2CH error estimates are different from each other and from 3CH error estimates, attribute this to biases, and work this idea out in section 7. But the differences may very well be a matter of representativeness. Both in 2CH and 3CH the "true" signal is the common signal shared by the two or three observation systems under consideration, and the "true" signal is determined by the system with the lowest resolution, see Stoffelen (1998). The authors should consider this.*

The reviewer raises issues related to *calibration* and *representativeness errors* in this comment.

Calibration: The calibration issue is closely related to the bias issue that we discuss. In the TC method, two of the data sets are calibrated against a third so that the three data sets are not biased with respect to each other. In the simplest form of the 3CH method, the data sets are not calibrated against each other, and thus some biases can exist. We investigate the effect of a 10% bias in one of the data sets and show even this fairly large bias does not greatly affect the 3CH results. However, even a much smaller bias affects the 2CH in a significant way. In addition, we also investigated the effect of biases in the related paper Anthes and Rieckh, 2018 by comparing the 3CH method to the triple co-location (TC) method (Stoffelen, 1998; Vogelzang et al., 2011). The results using the TC, in which the data sets were calibrated with respect to each other, were very similar to those using the 3CH method, confirming our results with the error model that biases do not cause large errors in the 3CH method.

To address this comment, we have added the following discussion to the beginning of Section 5:

“All four data sets have some degree of unknown bias for certain locations, altitudes, or atmospheric conditions; none of them represent the ultimate “truth” and there is no standard atmospheric data set for calibration. However, they have all been compared to other models or observations to one degree or another. We investigated the effect of biases in the related paper Anthes and Rieckh (2018) by comparing the 3CH method to the triple co-location (TC) method (Stoffelen, 1998; Vogelzang et al., 2011). The results using the TC method, in which the data sets were calibrated using the ERA-Interim data set as the calibration reference, were very similar to those using the 3CH method.”

2) *Representativeness errors*: We mention representativeness errors in the Introduction and discuss them in Section 5. The idealized model data sets with prescribed errors we developed are free of representativeness errors; they contain only known (specified) random and bias errors. However, a portion of the specified errors in the error model may be thought of as representativeness errors. Also, the 3CH method includes representativeness errors as part of the error estimates.

The two model (ERA-Interim and GFS) and two observational (RS and RO) data sets include random and bias errors. Representativeness errors, which are included in the error estimates from both the 2CH and 3CH method, may contribute to these random or bias errors. As our results show, random errors do not cause large differences between the 2CH and 3CH methods if the sample size is large enough; only bias errors do.

We showed that biases in one of the two data sets with respect to the other will cause large errors in the 2CH method. We believe this is the issue that Dr. Vogelzang discussed in his supplement discussion. We agree with his conclusion, that the 2CH method is very sensitive to calibration (bias) issues.

We added the following paragraph on page 16 to address the comment on representativeness errors:

“The two model and the RO data sets are representative of similar horizontal scales (~100 km), while the radiosonde data are in-situ point measurements and therefore represent a much smaller horizontal scale. However, many studies (e.g. Ho et al., 2010a,b; Kuo et al. 2004, Chen et al. 2011) have used radiosonde data as correlative data for verifying models, RO, and other data sets without applying corrections for representativeness errors. These results indicate that the different representative scales are not a significant source of error in the comparisons (unlike spatial and temporal sampling errors resulting from the time and spatial differences between the data sets, which we correct for). However, any representativeness errors are included in the error estimates using either the 2CH or 3CH method.”

3. *Another serious point is that the authors assume their error model to be valid without any further justification. A scatter plot of the data - the starting point of all analysis of data from multiple sources - would be helpful here. It will show if any calibration issues play a role and if errors indeed can be assumed to be independent of the observed value. The current presentation is more of a "trial and error" type. This should be considerably improved.*

   The error model is based on previous studies that have estimated the error profile of specific humidity. In order to justify the error model, we have added the following paragraph to the end of Section 3, page 5 line 19:

   "The error model is created based on error estimates of specific humidity from several studies (e.g. Kursinski, 1997; Collard and Healy, 2003; von Engeln and Nedoluha, 2005; Wang et al., 2013). For example, Collard and Healy (2003) found that, for tropical conditions, the percentage errors for RO specific humidity varied from approximately 10% near the surface to about 70% near 300 hPa. Other studies show the errors varying from about 10% near the surface to 100% in the upper troposphere (about 200 hPa). In our error model, we specified the STD to roughly approximate the STD of RO or RS data at Minamidaitojima, Japan (Anthes and Rieckh, 2018; Rieckh et al., 2018). The assumed STD of normalized q (percent error) given by Eq. (9) is consistent with the above empirical error estimates. Thus the error model is a reasonable one in terms of its magnitude and increase with height. Since it is intended to show the sensitivity of the 3CH and 2CH methods to varying degrees of correlation between two of the data sets used in the comparison, it is not necessary that the error model be a close replication of any particular observing system, just that the magnitude of the assumed errors and their vertical distribution be reasonable.

4. *A minor point: the link to Wriley (2003) gives no information on the history of the 3CH method.*

   Thank you; we have revised this sentence to:
   "W.J. Riley (2003), and references therein, provides a summary of the 3CH method."

Anonymous Referee #2

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-75, 2018.
Received and published: 27 March 2018

General comments

The authors compare error variances as generated from the three-cornered hat (3CH) and two-cornered hat (2CH) methods from simulated and real data sets. As expected, the 3CH results were less noisy and less sensitive to biases. The authors do a good job of presenting their widespread findings (from simulated and correlated data, as well as collocated real observations); however, I have some comments and suggestions that I would like to see addressed in their revision.

The 3CH assumes independent, uncorrelated observations, and its usefulness can be limited by *the sample size, as well as the variability the source data itself. The authors do mention several times that finite sample sizes will cause the cross-correlation terms to be non-zero. However, they do not seem to address the point that the error variances produced by the 3CH method (and likely 2CH as well) can be dominated by a data source that is largely different than the other two in the trio being analyzed. Thus, the size of the relative errors from the RO, GFS, ERA, and RS observations comes into play in the accuracy of the 3CH results in Section 6.*

Page 2 line 23 says that widely different errors associated with the three systems can reduce the accuracy of the estimates. We have inserted references to this in the revised paper. In our error model, the magnitudes of the errors in all three simulated data sets are similar. Previous studies of the errors of the systems tested in our paper (two model sets, ERA-Interim and GFS, and two observational data sets, RO and RS), indicate that the errors of all these systems are also of similar orders of magnitude (please see references in the paper and below). In addition, we recently used the 3CH to estimate the errors associated with these data sets and found that they were similar in magnitude (Anthes and Rieckh, 2018).

Some of the equation development is either incomplete or hard to follow. For example, on Pages 11-12, the authors derive equations for the estimated error variance of X,Y,Z. They start with the traditional equation for the variance (equations 2,3, and 4), but cross out some terms, then neglect the covariance term in the next step (because it's the 3CH estimate), then plug it back in. It took me a while to put it all together, so maybe the authors can add some additional text or format differently to help the reader along.

We agree and have revised the presentation of the equations by reducing the number of similar equations and clarifying the presentation. Please see also the response to Comment 1 of Reviewer 1.

The authors also seem to already possess knowledge regarding climatological processes and models to set up and normalize their model profile (section 3.1). Perhaps these numbers are taken from some of their previous work, but some additional text or references pertaining to the source or reasoning behind these numbers would be helpful.

We have provided considerably more discussion in the first part of Section 3 on the error model and its properties, including comparison with other studies of the estimated errors of specific humidity observations in the atmosphere. Please see our response to point 3 of Reviewer #1 (above) and the added paragraph discussing the error model.

Line edits
The grammar and sentence structure is quite good and easy to read, I only have a few line edits to offer.

Thank you for the careful reading. We have corrected all these typos.

Page 2
Line 11: Should be W.J. Riley, not W.J. Wriley

Corrected.

Line 13: 3CH, not 3HC
Corrected.

Page 8
Line 17: Should x,z be capitalized?

Section 3 had confusing notation with the upper and lower case variables. We have revised the text and equations to use only upper case X, Y and Z.

Page 9
Line 3: Should be There, not These
Corrected.

References (also included in revised paper)

Anthes, R.A. and Rieckh, T: Estimating observation and model error variances using multiple data sets. Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-487, 2018.

Chen, S.-Y., Huang, C.-Y., Kuo, Y.-H., and Sokolovskiy, S.: Observational Error Estimation of FORMOSAT-3/COSMIC GPS Radio Occultation Data, Mon. Wea. Rev., 139, 853–865, https://doi.org/10.1175/2010MWR3260.1, 2011.

Collard, A.D. and Healy, S.B.: The combined impact of future space-based atmospheric sounding instruments on numerical weather-prediction analysis fields: A simulation study, Q. J. R. Meteorol. Soc., 129, pp. 2741–2760 doi: 10.1256/qj.02.124, 2003.

Ho, Shu-Peng, Zhou, Xinjia, Kuo, Ying-Hwa, Hunt, D., and Wang, J-H.: Global evaluation of radiosonde water vapor systematic biases using GPS Radio Occultation from COSMIC and ECMWF analysis. Remote Sens., 2, 1320-1330; doi:10.3390/rs2051320, 2010a.

Ho, S.-P., Kuo, Y.-H., W. Schreiner, W. and Zhou, X.: Using SI-traceable Global Positioning System radio occultation measurements for climate monitoring, Bull. Am. Meteorol. Soc., 91(7), S36–S37, 2010b.

Kuo, Y.-H., Wee, T.-K., Sokolovskiy, S., Rocken, C., Schreiner, W., Hunt, D., and Anthes, R. A.: Inversion and error estimation of GPS radio occultation data, J. Meteor. Soc. Japan, 82, 507–531, 2004.

Kursinski, E. R., G. A. Hajj, G.A., Schofield, J.T., Linfield,R.P. and Hardy, K.R.: Observing Earth’s atmosphere with radio occultation measurements using the Global Positioning System, J. Geophys. Res., 102, 23,429–23,465, doi:10.1029/97JD01569, 1997.

Rieckh, T., Anthes, R.; Randel, W., Ho, S.-P.; and Foelsche, U.: Evaluating tropospheric humidity from GPS radio occultation, radiosonde, and AIRS from high-resolution time series. Atmos. Meas. Tech. Disc., https://doi.org/10.5194/amt-2017-486, 2018.

von Engeln, A. and Nedoluha, G.:Retrieval of temperature and water vapor profiles from radio occultation refractivity and bending angle measurements using an Optimal Estimation approach: a simulation study, Atmos. Chem. Phys., 5, 1665–1677, www.atmos-chem-phys.org/acp/5/1665/, 2005

Wang, B.R., Liu, X.-Y. NS WANG, J.-K.: Assessment of COSMIC radio occultation retrieval product using global radiosonde data. Atmos. Meas. Tech., 6, 1073–1083, www.atmos-meas-tech.net/6/1073/2013/ doi:10.5194/amt-6-1073-2013, 2013.

Evaluating two methods of estimating error variances from multiple data sets using an error model

Therese Rieckh[1,2] and Richard Anthes[1]

[1]COSMIC Program Office, University Corporation for Atmospheric Research, Colorado, U.S.A.

[2]Wegener Center for Climate and Global Change, University of Graz, Graz, Austria

Correspondence: Richard Anthes (anthes@ucar.edu)

Abstract. In this paper we compare two different methods of estimating the error variances of two or more independent data sets. One method, called the "three-cornered hat" (3CH) ~~or "triple co-location"~~ method, requires three data sets. Another method, which we call the "two-cornered hat" (2CH) method, requires only two data sets. Both methods assume that the errors of the data sets are not correlated~~and are unbiased~~. The 3CH method has been used in previous studies to estimate the error variances associated with a number of physical and geophysical data sets. Braun et al. (2001) used the ~~two-cornered hat (~~2CH ~~)~~ method to estimate the error variances associated with two observational data sets of total atmospheric water vapor.

~~In this paper we~~ We compare the 3CH and 2CH methods using a simple error model to simulate three and two data sets with various error correlations and biases. With this error model, we know the exact error variances and covariances, which we use to assess the accuracy of the 3CH and 2CH estimates. We examine the sensitivity of the estimated error variances to the degree of error correlation between two of the data sets as well as the sample size. We find that the 3CH method is less sensitive to these factors than the 2CH method and hence is more accurate. We also find that biases in one of the data sets has a minimal effect on the 3CH method, but can produce large errors in the 2CH method.

1 Introduction

In atmospheric sciences, observations and models are often combined with the goal of providing accurate and complete representations of the current or future state of the atmosphere. Knowing the error characteristics of observations and models is important to understanding the degree to which atmospheric phenomena of interest are accurately described and analyzed. Estimating observational and ~~modelling~~ modeling error characteristics are thus of inherent scientific interest. In addition, knowing the error characteristics are important for practical applications such as data assimilation and numerical weather prediction. In many modern data assimilation schemes, observations of a given type are weighted proportionally to the inverse of their error variance (e.g. Desroziers and Ivanov, 2001).

There are ~~at least three~~ several somewhat similar methods for estimating the error variances associated with two or more data sets. The "three-cornered hat" (3CH) ~~or~~ method and the closely related "triple co-location~~method" has~~" (TC) method have been used in physics, oceanography, and other scientific disciplines to estimate the errors associated with three independent data sets. Braun et al. (2001) combined two independent data sets, Global Positioning System (GPS) slant water vapor (SWV) and water vapor radiometer (WVR), to estimate the SWV and WVR errors. In analogy to the 3CH method, we refer to the Braun et al. (2001) method as the two-cornered hat (2CH) method. Kuo et al. (2004) and Chen et al. (2011) used the “apparent error method”, which is a variation of the 2CH method, to estimate the error of radio occultation (RO) observations using the known (or estimated) error variance of a forecast model.

The 3CH method was originally developed as the “N-cornered hat” method (Gray and Allan, 1974) for estimating error variances from ~~N~~ N atomic clocks. Requiring three data sets at a minimum, the 3CH method produces estimates of the error variances of each of three data sets. ~~W.J. Wriley (2003) provides~~ Riley (2003) , and references therein, provide a summary of the 3CH method~~and its history ()~~. Variations and enhancements of the method have been applied to many diverse geophysical data sets. The ~~3HC~~ 3CH method has been used to estimate the stability of GNSS clocks using the measured frequencies from multiple clocks (Ekstrom and Koppang, 2006; Griggs et al., 2014, 2015; Luna et al., 2017). Valty et al. (2013) used the 3CH method to estimate the geophysical load deformation computed from GRACE satellites, GPS vertical displacement measurements, and global general circulation (GCM) models. O’Carroll et al. (2008) compared three systems to measure sea-surface temperatures: two different radiometers and in situ observations from buoys. They discuss the assumption of neglecting the error correlations among the three data sets and the effect of representativeness errors. Anthes and Rieckh (2018) used the 3CH method to estimate the error variances of three observational (two radio occultation retrievals and radiosondes) and two model data sets using various combinations of the five data sets.

~~The 3CH~~ Stoffelen (1998) developed a closely related method, termed the “~~triple-co-location method” by Stoffelen (1998) in his study of estimated ocean~~ triple co-location” (TC) method to estimate errors of surface winds using three data sets~~, has~~ . Variations of Stoffelen’s TC method have been have been widely used in the ~~field~~ fields of oceanography and hydrometeorology ~~(Su et al., 2014; Gruber et al., 2016, e.g.) . O’Carroll et al. (2008) compared three systems to measure sea-surface temperatures: two different radiometers and in situ observations from buoys. They discuss the assumption of neglecting the error correlations among the three data sets and the effect of representativeness errors~~(e.g. Su et al., 2014; Gruber et al., 2016) . Roebeling et al. (2012) used the ~~triple co-location~~ TC method to estimate the errors associated with three ways of ~~estimating~~ measuring precipitation: the Spinning Enhanced Visible and Infrared Imager (SEVERI), weather radars, and ground-based rain gauges.

The major assumption in all of the above methods is that the errors of the three systems are uncorrelated~~and unbiased~~. Correlations between any or all of the three measurement systems will reduce the accuracy of the error estimates. Other factors that can reduce the accuracy include widely different errors associated with the three systems or a small sample size. These factors can lead to negative estimates of error variances, especially when the estimates are close to zero (Gray and Allan, 1974; Riley, 2003; Griggs

In this paper we estimate the effect of neglecting the error covariances using two or three simulated data sets for which the true error variances and covariances are known. We develop an error model to simulate the data sets with random and bias errors using a set of assumed True profiles. We then calculate the true error variance and covariance terms in the ~~three~~ simulated data sets and show the impact of neglecting these terms on the estimated error variances.

2 Error estimates using the 2CH and 3CH methods

We assume we have three data sets X, Y and Z that are all measuring the same physical variable, e.g. specific humidity, q, at the same location and time.

The *error variance* of the data set X is defined as

$$\mathrm{VAR_{err}}(\mathrm{X}) = ({}^{1}/{}_{n})\frac{1}{n}\sum(\mathrm{X}-\mathrm{True})^2 = (1/n)\frac{1}{n}\sum \mathrm{X}_{\mathrm{err}}^2 \tag{1}$$

where True is the true (but unknown) value of X (as well as Y and Z), $\mathrm{X_{err}}$ = (X – True), and n is the number of samples. O'Carroll et al. (2008) provide an excellent discussion of the meaning of "True" in the context of the 3CH method.

The *error covariance* of the data sets X and Y is defined as

$$\mathrm{COV_{err}}(\mathrm{X,Z}) = \frac{1}{n}\sum(\mathrm{X}-\mathrm{True})(\mathrm{Y}-\mathrm{True}) \tag{2}$$

In general, the errors of X, Y, and Z may be correlated or not.

2.1 Three-cornered hat (3CH) method

In the 3CH method, the relationship between the mean square error variances, the mean square (MS) differences of X and Y , MS(X – Y(MS(X–Y)), and their error variances and covariances are given by Eqs. (7)–(9) of Anthes and Rieckh (2018):

$$\begin{aligned}\mathrm{VAR_{err}}(\mathrm{X}) =& {}^{1}/{}_{2}{}^{1}/{}_{2}\,[\mathrm{MS(X-Y)+MS(X-Z)-MS(Y-Z)}]\\ &+\mathrm{COV_{err}(X,Y)}+\mathrm{COV_{err}(X,Z)}-\mathrm{COV_{err}(Y,Z)}\end{aligned} \tag{3}$$

$$\begin{aligned}\mathrm{VAR_{err}}(\mathrm{Y}) =& {}^{1}/{}_{2}{}^{1}/{}_{2}\,[\mathrm{MS(X-Y)+MS(Y-Z)-MS(X-Z)}]\\ &+\mathrm{COV_{err}(X,Y)}+\mathrm{COV_{err}(Y,Z)}-\mathrm{COV_{err}(X,Z)}\end{aligned} \tag{4}$$

$$\begin{aligned}\mathrm{VAR_{err}}(\mathrm{Z}) =& {}^{1}/{}_{2}{}^{1}/{}_{2}\,[\mathrm{MS(X-Z)+MS(Y-Z)-MS(X-Y)}]\\ &+\mathrm{COV_{err}(X,Z)}+\mathrm{COV_{err}(Y,Z)}-\mathrm{COV_{err}(X,Y)}\end{aligned} \tag{5}$$

We note that Eqs. (3)–(5) are identical to Eq. (1) in O'Carroll et al. (2008) in their application of the TC method to estimate the error variances of sea surface temperature using observations from two satellite systems and in-situ buoys.

The last three covariance terms in Eqs. ??(3)–?? are the terms that we neglect when using real data to estimate (5) are unknown for real data sets and are neglected when estimating the error variances of the real data sets X, Y and Z. This assumption is valid for no correlation between the errors of the data sets and a large sample size. The estimated error variances are therefore:

$$\mathrm{VAR_{err}(X)_{est}} = {}^{1}/_{2}\,[\mathrm{MS(X-Y)} + \mathrm{MS(X-Z)} - \mathrm{MS(Y-Z)}] \tag{3a}$$

$$\mathrm{VAR_{err}(Y)_{est}} = {}^{1}/_{2}\,[\mathrm{MS(X-Y)} + \mathrm{MS(Y-Z)} - \mathrm{MS(X-Z)}] \tag{4a}$$

$$\mathrm{VAR_{err}(Z)_{est}} = {}^{1}/_{2}\,[\mathrm{MS(X-Z)} + \mathrm{MS(Y-Z)} - \mathrm{MS(X-Y)}] \tag{5a}$$

2.2 Two-cornered hat (2CH) method

In the The 2CH method , there are uses only two data sets, X and Z. To derive the relationship between the error variances of The derivation is presented in Appendix A. The 2CH method error variances for X and Z given the sums and differences of the two data sets, we first write $\mathrm{X = (True + X_{err})}$ and $\mathrm{Z = (True + Z_{err})}$, then add X and Z and square the sum: $\mathrm{(X+Z)^2 = 4\,True^2 + 4\,True\,(X_{err} + Z_{err}) + (X_{err} + Z_{err})^2}$ Equation A3 is summed over all the data pairs to get

$$\sum \mathrm{(X+Z)^2} = 4\sum \mathrm{True^2} + \sum \left[\mathrm{(X_{err} + Z_{err})^2 + 4True\,(X_{err} + Z_{err})}\right]$$

$$\begin{aligned}\mathrm{MS(X+Z)} =\, & 4\mathrm{MS(True)} + \mathrm{VAR_{err}(X)} + \mathrm{VAR_{err}(Z)} \\ & + 2\mathrm{COV_{err}(X,Z)} + 4\mathrm{M(True, X_{err} + Z_{err})}\end{aligned}$$

where

$\mathrm{M(X)} = ({}^{1}/_{n})\sum \mathrm{X}$

$\mathrm{M(X,Y)} = \mathrm{M(X\cdot Y)} = ({}^{1}/_{n})\sum \mathrm{X\,Z}$.We then subtract Z from X and square the difference to get

$$\mathrm{MS(X-Z)} = \mathrm{VAR_{err}(X)} + \mathrm{VAR_{err}(Z)} - 2\mathrm{COV_{err}(X,Z)}$$

Finally, we solve for MS(True) by subtracting Eq. A7 from Eq. A6

$$4\mathrm{MS(True)} = \mathrm{MS(X+Z)} - \mathrm{MS(X-Z)} - 4\mathrm{COV_{err}(X,Z)} - 4\mathrm{M(True, X_{err} + Z_{err})}$$

By squaring the expression $\mathrm{X = (True + X_{err})}$ we get the exact expression for the $\mathrm{VAR_{err}(X)}$

$$\mathrm{VAR_{err}(X)} = \mathrm{MS(X)} - \mathrm{MS(True)} - 2\mathrm{M(True, X_{err})}$$

and substituting for MS(True) from Eq. A8 gives are given as

$$\begin{aligned}\mathrm{VAR_{err}(X)} =\, & \mathrm{MS(X)} - {}^{1}/_{4}\,[\mathrm{MS(X+Z)} - \mathrm{MS(X-Z)}] \\ & - 2\mathrm{M(True, X_{err})} + \mathrm{COV_{err}(X,Z)} + \mathrm{M(True, X_{err} + Z_{err})}\end{aligned} \tag{6}$$

and

$$\begin{aligned}\mathrm{VAR_{err}(XZ)} =\, & \mathrm{MS(X) - MS(X+Z) - MS(X-Z)}/4\mathrm{MS(Z)} - {}^{1}/_{4}\,[\mathrm{MS(X+Z)} - \mathrm{MS(X-Z)}] \\ & -2-2\mathrm{M(True, XTrue, Z_{err})} + \mathrm{COV_{err}(X,Z)} + \mathrm{M(True, X_{err} + Z_{err})}.\end{aligned} \tag{7}$$

Omitting the last three error terms, we obtain:

$$\mathrm{VAR_{err}(X)_{est} = MS(X) - [MS(X+Z) - MS(X-Z)]/4} \quad (11a)$$

Similarly, we obtain

$$\mathrm{VAR_{err}(Z) = MS(Z) - [MS(X+Z) - MS(X-Z)]/4 - 2M(True, Z_{err}) + COV_{err}(X,Z) + M(True, X_{err} + Z_{err})}$$

and again neglecting the where M(X)is the mean value of X. The last three error terms we obtain

$$\mathrm{VAR_{err}(Z)_{est} = MS(Z) - [MS(X+Z) - MS(X-Z)]/4} \quad (12a)$$

Equations 11a and 12a are given as

$$\mathrm{VAR_{err}(X) = MS(X) - {}^{1}/_{4}\,[MS(X+Z) - MS(X-Z)] - 2M(True, X_{err}) + COV_{err}(X,Z) + M(True, X_{err} + Z_{err})} \quad (8)$$

and

$$\mathrm{VAR_{err}(Z) = MS(Z) - {}^{1}/_{4}\,[MS(X+Z) - MS(X-Z)] - 2M(True, Z_{err}) + COV_{err}(X,Z) + M(True, X_{err} + Z_{err}).} \quad (9)$$

The last three error terms in Eq. (8) and (9) are unknown for real data sets and are neglected when estimating the error variances of the real data sets X and Z:

$$\mathrm{VAR_{err}(X)_{est} = MS(X) - {}^{1}/_{4}\,[MS(X+Z) - MS(X-Z)]} \quad (6a)$$

$$\mathrm{VAR_{err}(Z)_{est} = MS(Z) - {}^{1}/_{4}\,[MS(X+Z) - MS(X-Z)].} \quad (7a)$$

Equations (6a) and (7a) are equivalent to Eqs. (12) and (13) of Braun et al. (2001).

2.3 Side note: apparent error method

We note that Eq. A7 is the basis for the "apparent error" (AE) method in which X is an observed data set (X_{obs}) and Z is a forecast of the X data set (X_{fcst})

$\mathrm{AE = X_{obs} - X_{fcst}}$

$\mathrm{VAR_{err}(X_{obs}) = MS(X_{obs} - X_{fcst}) - VAR_{err}(X_{fcst}) + 2COV_{err}(X_{obs}, X_{fcst})}$ The apparent error is equivalent to the Observation minus Background (O – B) statistic used in data assimilation studies. In the apparent error method, the correlation of errors between the observations and forecasts is assumed negligible and the error variance of the forecast is obtained from an independent estimate.

2.3 Comparison of neglected terms in 3CH and 2CH methods

In the 3CH method, the neglected error terms when computing $\mathrm{VAR_{err}(X)}$ with Eq. ?? (3) are

$\mathrm{COV_{err}(X, Y) + COV_{err}(X, Z) - COV_{err}(Y, Z)}$.

In the 2CH method, the neglected error terms when computing $\mathrm{VAR_{err}}$ with Eq. A10 (8) are

$\mathrm{-\,2\,M(True, X_{err}) + COV_{err}(X, Z) + M(True, X_{err} + Z_{err})}$.

We note that the neglected error terms in the 2CH method contain terms involving the product of True with errors, unlike in the 3CH method. Because True is typically an order of magnitude greater than the errors, these terms are likely much larger than the neglected terms involving only products of errors, as in the 3CH method. We also note that if the X and Z errors are random and uncorrelated, all of the error terms will be zero for an infinite sample size. However, for finite sample sizes, these terms will be non-zero even if these conditions are met.

2.4 Side note: apparent error method

We note that Eq. (A7) in Appendix A is the basis for the "apparent error" (AE) method in which X is an observed data set ($\mathrm{X_{obs}}$) and Z is a forecast of the X data set ($\mathrm{X_{fcst}}$)

$\mathrm{AE = X_{obs} - X_{fcst}}$

$\mathrm{VAR_{err}(X_{obs}) = MS(X_{obs} - X_{fcst}) - VAR_{err}(X_{fcst}) + 2\,COV_{err}(X_{obs}, X_{fcst})}$

The apparent error is equivalent to the Observation minus Background (O – B) statistic used in data assimilation studies. In the apparent error method, the correlation of errors between the observations and forecasts is assumed negligible and the error variance of the forecast is obtained from an independent estimate.

3 Generation of True data set and three simulated data sets with errors

We first generate a set of n vertical profiles of a variable, True, which we take as specific humidity from the ERA-Interim reanalysis (Dee et al., 2011), normalized by the mean specific humidity $\bar{q}$ averaged over all samples. We next generate three data sets X, Y and Z that are approximations of True (True plus random errors), where the errors of X and Z are correlated to a degree that we can control. For simplicity, we assume the errors for Y are always uncorrelated with those of X and Z. This is analogous to a system of three observational systems in which the errors of two of them are correlated, but the errors of the third are not. In this section, only random errors are considered. The effect of bias errors is discussed in Sect. 4.4.2.

We then look at the magnitude of the error terms in the 2CH and 3CH methods with various assumed correlation coefficients between the errors in X and Z and compare the estimated error variances of X, Y and Z with their true error variances. Our tests will show the impact of the neglect of the error terms depending on the degree of error correlation between X and Z.

3.1 Model error profile

~~In Sections 3–5 we assume the mean error (bias) is zero for all the modeled data sets X, Y and Z. The effect of adding a bias error to Z is considered in Sect. ??. The assumed~~ The assumed normalized random error model for X is given by

$$\mathrm{X_{err}} = \mathrm{random}[-1.7, 1.7] \cdot \sout{\mathrm{CLIMO}} \cdot \mathrm{STD} \tag{10}$$

$$\mathrm{STD} = 0.1 + 0.00042 \cdot (1000 - p) \tag{11}$$

where random[–1.7, 1.7] is a random number between –1.7 and 1.7, ~~CLIMO is the average of True over all n samples (called CLIMO in analogy to using a climatological value of q),~~ and the standard deviation STD increases linearly with decreasing pressure p from a value of 0.1 (10 %) at 1000 hPa to a value of 0.436 (43.6 %) at 200 hPa. In the estimated error calculations below, ~~we normalize~~ the normalized $\mathrm{X_{err}}$ ~~by CLIMO so that it~~ is expressed as a % error and the variance is expressed as $\%^2$.

The error model is created based on error estimates of specific humidity from several studies (e.g. Kursinski et al., 1997; Collard and Hea For example, Collard and Healy (2003) found that, for tropical conditions, the percentage errors for RO specific humidity varied from approximately 10 % near the surface to about 70 % near 300 hPa. Other studies show the errors varying from about 10 % near the surface to 100 % in the upper troposphere (about 200 hPa). In our error model, we specified the STD to roughly approximate the STD of RO or RS data at Minamidaitojima, Japan (Anthes and Rieckh, 2018; Rieckh et al., 2018) . The assumed STD of normalized q (percent error) given by Eq. (11) is consistent with the above empirical error estimates. Thus the error model is a reasonable one in terms of its magnitude and increase with height. Since it is intended to show the sensitivity of the 3CH and 2CH methods to varying degrees of correlation between two of the data sets used in the comparison, it is not necessary that the error model be a close replication of any particular observing system, just that the magnitude of the assumed errors and their vertical distribution be reasonable.

3.1 Calculation of correlated errors

We first generate the random error profiles $\mathrm{X_{err}}$, $\mathrm{Y_{err}}$, and $\mathrm{Q_{err}}$. All of these error profiles are uncorrelated. In general all three error profiles $\mathrm{X_{err}}$, $\mathrm{Y_{err}}$, and $\mathrm{Q_{err}}$ may have different standard deviations, but for these tests we assume all standard deviations vary according to Eq. ~~11~~ (11) for simplicity. We now generate the error profile $\mathrm{Z_{err}}$ as a linear combination of $\mathrm{X_{err}}$ and $\mathrm{Q_{err}}$:

$$\mathrm{Z_{err}} = \sout{(a \cdot \mathrm{X_{err}} + \mathrm{Q_{err}})/(1+a)} \, \frac{a \cdot \mathrm{X_{err}} + \mathrm{Q_{err}}}{1+a} \tag{12}$$

where a is a specified constant parameter that determines the degree of correlation between $\mathrm{Z_{err}}$ and $\mathrm{X_{err}}$. If $a = 0$, $\mathrm{Z_{err}} = \mathrm{Q_{err}}$ and the errors of X, Y and Z are all uncorrelated. If a is not equal 0, the errors of Z are correlated with the errors of X. ~~In this simple model, the errors of Zand Y and the errors of X and Y are always uncorrelated.~~ Figure 1a shows one set of simulated error profiles for the unnormalized X, Q, and several Z for different values of a. The transition of $\mathrm{Z_{err}}$ from $\mathrm{Q_{err}}$ to $\mathrm{X_{err}}$ can be seen easily in the black box around 600 hPa. $\mathrm{Q_{err}}$ (bright green line) is positive (around 1 g kg^{-1}) and $\mathrm{Z_{err}}$ for $a = 0$ is identical to that. For $a = 0.1$, $\mathrm{Z_{err}}$ has a slightly smaller positive value, and $\mathrm{Z_{err}}$ becomes negative as a increases, becoming practically identical with $\mathrm{X_{err}}$ (dotted line) for $a = 100$.

Figure 1b shows the mean and standard deviation for the unnormalized Z specific humidity profiles for $a = 0$. Since the $\mathrm{Z_{err}}$ are created by combining the two random errors $\mathrm{Q_{err}}$ and $\mathrm{X_{err}}$, the $\mathrm{Z_{err}}$ are overall closer to zero (as can be seen in Fig. 1a in

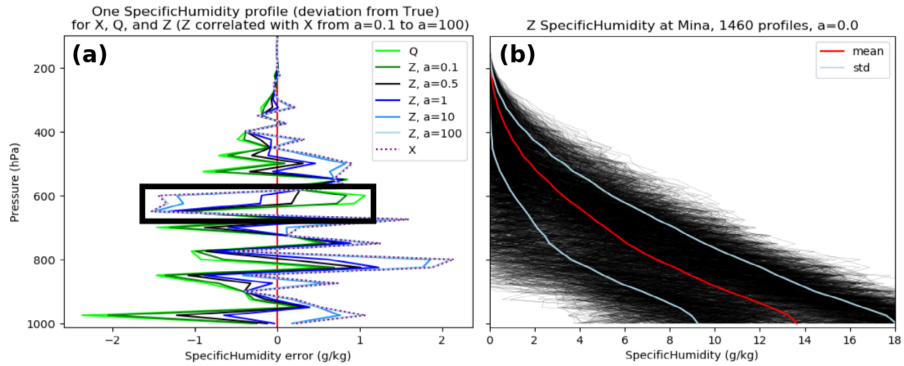

Figure 1. (a) One set of error profiles with the unnormalized X_{err} (dotted line), Q_{err} (solid lime green line) and Z_{err} (various solid lines for different values of correlation parameter a). Q_{err} is uncorrelated with X_{err}. For $a = 0$, $Z_{err} = Q_{err}$ and Z_{err} and X_{err} are uncorrelated. As a increases, Z_{err} is increasingly correlated with X_{err} and looks less like Q_{err} and more like X_{err}. For very large a ($a = 100$), Z_{err} (light blue solid line) is almost equal to X_{err} (dotted line). (b) 1460 profiles of the unnormalized data set Z with zero correlation with X ($a = 0$).

the black box.) This will result in a smaller standard deviation of Z if $0 < a < \infty$ (especially for values of a close to 1), and will also decrease the true error variance of Z.

3.2 Relationship of correlation, error variance, and error covariance terms

The correlation coefficient r between the Xand Z errors, r_{xz} also denoted by r(X_{err} , and Z_{err}), is given by

$$r_{xz} = (^1/_n)\frac{1}{n}\sum X_{err} - (X_{err})_m Z_{err} - (Z_{err})_m / \sigma_x\sigma_z \frac{(X_{err} - \overline{X}_{err})(Z_{err} - \overline{Z}_{err})}{\sigma_{X_{err}}\sigma_{Z_{err}}} \tag{13}$$

where $(X_{err})_m$ and $(Z_{err})_m$ $\overline{X}_{err}$ and $\overline{Z}_{err}$ are the mean values of the Xand Z errors X_{err} and Z_{err} respectively (zero in our error model) and σ_x and σ_z are $\sigma_{X_{err}}$ and $\sigma_{Z_{err}}$ are the standard deviations of X_{err} and Z_{err} given by

$\sigma_x^2 = (^1/_n)\sum X_{err}^2$ $\sigma_{X_{err}} = \sqrt{\frac{1}{n}\sum X_{err}^2} \equiv (VAR_{err}(X))^{1/2}$ using Eq. (1)

$\sigma_z^2 = (^1/_n)\sum Z_{err}^2$ $\sigma_{Z_{err}} = \sqrt{\frac{1}{n}\sum Z_{err}^2} \equiv (VAR_{err}(Z))^{1/2}$ using Eq. (1)

The correlation coefficient

The relationship between the correlation coefficient and covariance between X_{err} and Z_{err} may also be written as is therefore

r$COV(X_{err}, Z_{err}) = r\,\sigma_{X_{err}}\sigma_{Z_{err}} \equiv COV_{err}(X, Z)$ $/\,\sigma(X_{err})\,\sigma(Z_{err})$ using Eq. (2).

where $\sigma(X_{err})$ and $\sigma(Z_{err})$ are the standard deviations of the X and Z errors respectively.

$$COV_{err}(X, Z) = r_{xz}\sigma_x\sigma_z$$

where x and z are the normalized errors associated with X and Z respectively.

It can be shown that for this error model $r_{xz} = a/(1+a^2)^{1/2}$ $\sigma_z = (1+a^2)^{1/2}/(1+a)\sigma_x \sigma_z^2 = (1+a^2)/(1+a)^2\sigma_x^2$ $VAR_{err}(Z) = (1+a^2)/(1+a)^2 VAR_{err}(X)$ $COV_{err}(X, Z) = a/(1+a) VAR_{err}(X)$ Thus the particular error model

$$r = \frac{a}{(1+a^2)^{1/2}} \tag{14}$$

$$VAR_{err}(Z) = \frac{(1+a^2)}{(1+a)^2} VAR_{err}(X) \tag{15}$$

$$COV_{err}(X, Z) = \frac{a}{(1+a)} VAR_{err}(X) \tag{16}$$

The correlation between X and Z errors can be varied according to Eq. (14) by varying the parameter a, as shown in Table 1. In the Table 1 example, the STD for the normalized XNote that VAR_{err} is assumed to be a constant 10($VAR = 100^2$) rather than varying the STD according to Eq. 11. Note that $\sigma_z \leq \sigma_x$ and (Z) $\leq$ and $VAR_{err}(X)$ and that the maximum difference between $VAR_{err\,err}(X)$ and $VAR_{err\,err}(Z)$ occurs for a = 1 when σ_z is $VAR_{err}(Z) = 0.5\, VAR_{err}(X)$ ($\sigma_{Z_{err}}$ = 0.707 σ_x or $\sqrt{2}/2 \cdot \sigma_x \sigma_{X_{err}}$).

Table 1. Relationship between normalized error variances and standard deviations of data sets X and Z for different values of a($x \equiv$ normalized X_{err} and $z \equiv$ normalized Z_{err}).

a	r_{xz}	σ_z/σ_x $\sigma_{Z_{err}}/\sigma_{X_{err}}$	$VAR_{err}(Z)$ ($\%^2$)	$COV_{err}(X, Z)$ ($\%^2$)
0.0	0.0	1.0	1.00 $VAR_{err}(X)$	0.0
0.1	0.0995	0.9136	0.84 $VAR_{err}(X)$	0.0909 $VAR_{err}(X)$
0.3	0.287	0.803	0.65 $VAR_{err}(X)$	0.230 $VAR_{err}(X)$
0.5	0.447	0.745	0.55 $VAR_{err}(X)$	0.333 $VAR_{err}(X)$
1.0	0.707	0.707	0.50 $VAR_{err}(X)$	0.500 $VAR_{err}(X)$
2.0	0.894	0.745	0.55 $VAR_{err}(X)$	0.666 $VAR_{err}(X)$
10.0	0.995	0.913	0.83 $VAR_{err}(X)$	0.908 $VAR_{err}(X)$
100.0	0.99995	0.9901	0.98 $VAR_{err}(X)$	0.990 $VAR_{err}(X)$
∞	1.0	1.0	1.00 $VAR_{err}(X)$	1.0 $VAR_{err}(X)$

3.3 Summary of generation of True data set and simulated data sets with random and systematic errors

- We use 2007 ERA-Interim specific humidity q profiles for a location near Minamidaitojima (hereafter Mina), Japan, which is located on Okinawa at 25.6°N 131.5°W. These There are four model data profiles per day or n = 1460 profiles.
- Assume each vertical profile of q has no error, and normalize the q at each level by the sample mean ($\bar{q}$) at that level. This is the True data set.
- Generate three different and independent random error profiles X_{err}, Y_{err} and Q_{err} from Eqs. 10 and 11(10) and (11).
- Generate Z_{err} from Eq. 12 for various specified values of a to control correlation using Eq. (12).

– Add X_{err}, Y_{err} and Z_{err} to True to obtain ~~the 1460 simulated profiles~~ X, Y and Z~~respectively~~, the 1460 simulated normalized profiles of q.

– Compute the normalized estimated error variance profiles of X, Y and Z for the 3CH and 2CH methods according to Eqs. ~~**??**, **??**, and **??** neglecting the COV terms~~(3a)–(7a) (which have neglected all COV terms) and compare with the true error variances, which can be computed exactly from the full Eqs. ~~**??**~~(3)–~~**??**~~ (9) including the known values of the covariance terms. ~~All variance and covariance terms are normalized by the 2007 average value of specific humidity (CLIMO), computed from the True data set.~~ A value of ~~$a = 0$~~ $a = 0$ should give the most accurate estimation of error variances because all covariance terms will be close to zero (they won't be exactly zero because the sample size n is finite).

~~4 Example of simulated data profiles with random errors added~~

~~Figure 1a shows one set of simulated error profiles for X, Q, and several Z for different values of a. The transition of Z_{err} from Q_{err} to X_{err} can be seen easily in the black box around 600hPa. Q_{err} (bright green line) is positive (around 1gkg^{-1}) and Z_{err} for $a = 0$ is identical to that. For $a = 0.1$, Z_{err} has a slightly smaller positive value, and Z_{err} becomes negative as a increases, becoming practically identical with X_{err} (dotted line) for $a = 100$.~~

~~Figure 1b shows the mean and standard deviation for the Z specific humidity profiles for $a = 0$. Since the Z_{err} are created by combining the two random errors Q_{err} and X_{err}, the Z_{err} are overall closer to zero (as can be seen in Fig. 1 1a in the black box.) This will result in a smaller standard deviation of Z if $0 < a < \infty$ (especially for values of a close to 1), and will also decrease the true error variance of Z.~~

~~(a) One set of error profiles with X_{err} (dotted line), Q_{err} (solid lime green line) and Z_{err} (various solid lines for different values of a). Q_{err} is uncorrelated with X_{err}. For $a = 0$, $Z_{err} = Q_{err}$ and Z_{err} and X_{err} are uncorrelated. As a increases, Z_{err} is increasingly correlated with X_{err} and looks less like Q_{err} and more like X_{err}. For very large a ($a = 100$), Z_{err} (light blue solid line) is almost equal to X_{err} (dotted line). (b) 1460 profiles of data set Z with zero correlation with X ($a = 0$).~~

4 Effect of error correlations on estimated error variances

We now derive expressions for the estimated values of the error variances ~~and their standard deviations~~ for X, Y and Z for this error model and show how the correlations between X_{err} and Z_{err} affect the approximate values using the 3CH and 2CH methods. This will give some insight into how correlations between actual observed data sets will affect estimates of their error variances and standard deviations. To make results more readily comparable to previous studies, instead of showing the error variance, we show the square root of the error variance, or the error standard deviation, in most figures.

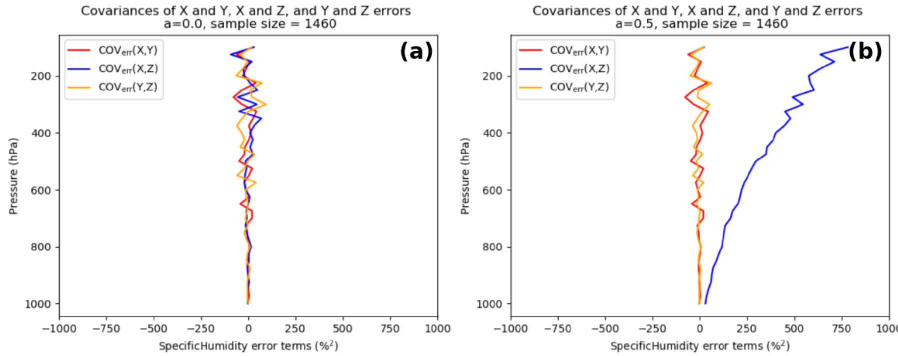

Figure 2. Vertical profiles of normalized error covariances (X, Y), (Y, Z) and (X, Z) for: (a) a = 0 and (b) a = 0.5.

4.1 Effect of error correlations on 3CH method

In the 3CH method, three covariance terms are neglected. ~~For uncorrelated errors between data sets, these~~ These terms are zero ~~(~~for an infinite sample size ~~)~~for uncorrelated errors between the data sets. Correlation between the errors of two data sets will lead to a non-zero covariance term, which becomes larger for larger correlations. The error covariances $\mathrm{COV_{err}}$(X, Z), $\mathrm{COV_{err}}$(X, Y) and $\mathrm{COV_{err}}$(Y, Z) are shown in Fig. 2 for $a = 0$ and 0.5. Note that ~~the error covariances of X and Y (red line) and Y and Z (orange line)~~ all error covariances oscillate around zero (and are zero for an infinitely large sample size) for $a = 0$ (Fig. 2a). For $a = 0.5$, the $\mathrm{COV_{err}}$(X, Z) term increases with decreasing pressure, reaching a magnitude of about 800 $\%^2$ at 100 hPa (Fig. 2b).

~~Completing the derivation of the error variances now, Eq. ?? becomes~~

$$\mathrm{VAR_{err}(X)} = {}^1/_2[\mathrm{MS(X-Y) + MS(X-Z) - MS(Y-Z)}] + \mathrm{COV_{err}(X,Y)} + \mathrm{COV_{err}(X,Z)} - \mathrm{COV_{err}(Y,Z)}$$

~~where the terms crossed out are zero (for an infinite data set) because Y_{err} is uncorrelated with X_{err} and Z_{err} in our error model. In the three cornered hat method, the estimate of VAR~~For our model (error correlation of X and Z) and an infinite sample size, $\mathrm{COV_{err}}$(X~~) is obtained by neglecting the term $\mathrm{COV_{err}}$(X, Z)~~, Z) is the only non-zero neglected term in Eq. (3a). Therefore Eq. (3a) can also be expressed as

$$\mathrm{VAR_{err}(X)_{est}} = {}^1/_2\mathrm{MS(X-Y) + MS(X-Z) - MS(Y-Z)}\,\mathrm{VAR_{err}(X) - COV_{err}(X,Z)} \tag{2b}$$

~~or~~

$$\mathrm{VAR_{err}(X)_{est}} = \mathrm{VAR_{err}(X) - COV_{err}(X,Z)} \tag{2b}$$

Using Eq. ~~16~~ (16) for this error model yields

$$\mathrm{VAR_{err}(X)_{est}} = [1/(1+a)]\mathrm{VAR_{err}(X)}. \tag{2c}$$

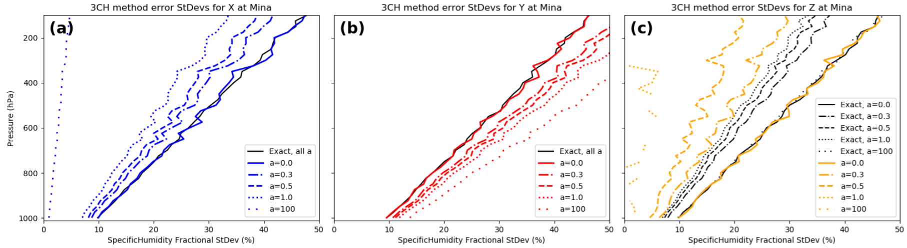

Figure 3. (a) Estimated standard deviation of X_{err} for values of a = 0, 0.3, 0.5, 1.0 and 100 computed from Eq. (3a)–(5a). Exact STD computed from data set X (solid black profile). (b) Same as (a) except for estimated STD of Y_{err}. (c) Same as (a) except for estimated STD of Z_{err}. Note that for different values of a, the exact STD of X_{err} and Y_{err} are always the same. The exact STD of decreases Z_{err} as a increases for $0 < a < 1.0$. For $a > 1.0$, the exact STD of Z_{err} increases as a increases, becoming equal to the STD of X_{err} for $a = \infty$ due to the way our error model is defined (see also Table 1). For $a = 100$ the correlation between X and Z errors is 0.99995 (Table 1; black dotted line for $a = 100$ almost identical to black solid line for $a = 0$).

Hence for $a > 0$ the estimated error variance of X is always less than the true value, as seen in Fig. 3a.

We next consider the effect of the X and Z error correlation on the estimate for Y error variance. Crossing out the zero termsFor our error model, Eq. ?? becomes

$$\mathrm{VAR_{err}(Y)} = {}^{1}/_{2}[\mathrm{MS(X-Y)} + \mathrm{MS(Y-Z)} - \mathrm{MS(X-Z)}] + \mathrm{COV_{err}(X,Y)} + \mathrm{COV_{err}(Y,Z)} - \mathrm{COV_{err}(X,Z)}$$

or

$$\mathrm{VAR_{err}(Y)_{est}} = \mathrm{MS(X-Y)} + \mathrm{MS(Y-Z)} - \mathrm{MS(X-Z)} \tag{3a}$$

(4a) can be expressed as

$$\mathrm{VAR_{err}(Y)_{est}} = \mathrm{VAR_{err}(Y)} + \mathrm{COV_{err}(X,Z)}. \tag{3b}$$

Substituting for the COV term from Eq. 16 (16) and noting that, for our error model, $\mathrm{VAR_{err}(X)} = \mathrm{VAR_{err}(Y)}$ we obtain

$$\mathrm{VAR_{err}(Y)_{est}} = [(1+2a)/(1+a)]\mathrm{VAR_{err}(Y)} \tag{3c}$$

Thus the estimated error variance for Y is always greater that the true value for $a > 0$$a > 0$, which is seen in Fig. 3b.

Lastly, we consider the effect of the X and Z correlation on the estimate for the Z error variance. Eq. ??, with the zero terms for this error model crossed out, becomes

$$\mathrm{VAR_{err}(Z)} = {}^{1}/_{2}[\mathrm{MS(X-Z)} + \mathrm{MS(Y-Z)} - \mathrm{MS(X-Y)}] + \mathrm{COV_{err}(X,Z)} + \mathrm{COV_{err}(Y,Z)} - \mathrm{COV_{err}(X,Y)}$$

or

$$\mathrm{VAR_{err}(Z)_{est} = MS(X-Z) + MS(Y-Z) - MS(X-Y)} \quad (4a)$$

(5a) can be expressed as

$$\mathrm{VAR_{err}(Z)_{est} = VAR_{err}(Z) - COV_{err}(X, Z)}. \quad (4b)$$

Substituting for the COV term from Eq. 16 (16) and using Eq. 15 we obtain (15) we obtain

$$\mathrm{VAR_{err}(Z)_{est}} = [(1-a)/(1+a^2)]\mathrm{VAR_{err}(Z)} \quad (4c)$$

so that the estimated error variance for Z is less than the true value for a >0, which is illustrated in Fig. 3c. Finally, for a >1, Eq. ?? (4c) shows that the estimated error variance of Z is negative and the STD is undefined. For a = 1.0, the estimated error variance of Z is zero for an infinite data set, but oscillates around zero because of our finite data set. Thus the estimated STD of VAR$\mathrm{Z_{err}}$ is undefined at some levels (Fig. 3c).

(a) Estimated standard deviation of $\mathrm{X_{err}}$ for values of a = 0, 0.3, 0.5, 1.0 and 100 computed from Eq. ??. Exact STD computed from data set X (solid black profile). (b) Same as (a) except for estimated STD of $\mathrm{Y_{err}}$. (c) Same as (a) except for estimated STD of $\mathrm{Z_{err}}$. Note that for different values of a, the exact STD of $\mathrm{X_{err}}$ and $\mathrm{Y_{err}}$ are always the same. The exact STD of decreases $\mathrm{Z_{err}}$ as a increases for $0 < a < 1.0$. For $a > 1.0$, the exact STD of $\mathrm{Z_{err}}$ increases as a increases, becoming equal to the STD of $\mathrm{X_{err}}$ for $a = \infty$ due to the way our error model is defined (see also Table 1). For $a = 100$ the correlation between X and Z errors is 0.99995 (Table 1; black dotted line for $a = 100$ almost identical to black solid line for $a = 0$).

4.2 Summary of error correlations on 3CH method

The true values of STD($\mathrm{X_{err}}$) and STD($\mathrm{Y_{err}}$) are always the same, calculated from Eq. 11. The true values of STD($\mathrm{Z_{err}}$) are similar to STD($\mathrm{X_{err}}$) for $a = 0$ and equal to STD($\mathrm{X_{err}}$) for $a = \infty$. For $0 < a < \infty$ STD($\mathrm{Z_{err}}$) is less than STD($\mathrm{X_{err}}$) and reaches a minimum of 0.707 STD($\mathrm{X_{err}}$) for $a = 1$ (Table 1).

The COVFor the 3CH method, correlation between the data sets X and Z has the following affect on their computed error variances (when covariance terms are neglected):

VAR$_{\mathrm{err}}$(X, Z), which is neglected in all of the approximate calculations, varies from 0 for $a = 0$ to)$_{\mathrm{est}}$ < VAR$_{\mathrm{err}}$(X)for large a. Thus for large a, from Eq. 2b we find that the estimated $_{\mathrm{true}}$

VAR$_{\mathrm{err}}$(X)tends to zero as a increases.

For large a, from Eq. ?? we see that the estimated value of Y)$_{\mathrm{est}}$ > VAR$_{\mathrm{err}}$(Y)tends toward 2 $_{\mathrm{true}}$

VAR$_{\mathrm{err}}$(X). The approximate STD then is therefore $\sqrt{2}$ times the true STD for large a, and this is seen in the plot for $a = 100$ (Fig. 3b). Z)$_{\mathrm{est}}$ < VAR$_{\mathrm{err}}$(Z)$_{\mathrm{true}}$

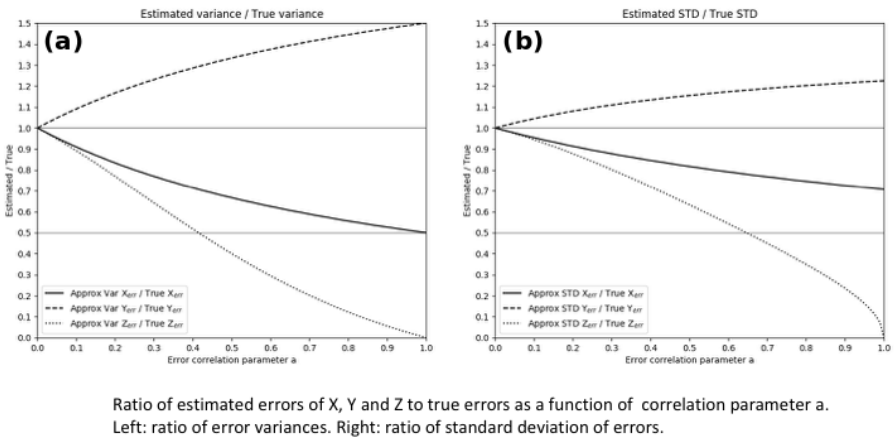

Figure 4. Ratio of estimated (a) error variances to true variances and (b) estimated STD to true STD for the 3CH method for values of error correlation parameter a ranging from 0 to 1.0.

Figure 4 shows the ratios of the approximate error variances and standard deviations to the true values for a ranging from 0 to 1 for the 3CH method. As the correlation parameter a increases from 0, the ~~ratios of the approximate to true errors increases~~differences between the estimated and true error variances increase. For a modest value of $a = 0.2$ (correlation coefficient between X and Z errors of 0.196), the errors in the STD estimates are –9 % for X, +8 % for Y and –12 % for Z. As the correlation between the X and Z error reach 0.5, the percentage errors for the X, Y and Z estimates reach –18 %, +14.5 % and –37 % respectively. To the extent that this error model gives an idea of the effect of the correlation between the errors of two of the three data sets, estimates of error standard deviations using a large sample of real data should be accurate to within approximately 10 % for correlation coefficients between data errors of 0.2 or below, and within 25 % for correlation coefficients of around 0.3. The effect of the correlation between Z and X errors on the estimated error variance is greatest on the estimated Z error variance.

4.3 Effect of error correlations on 2CH method

We next examine how error correlations in our error model affect the 2CH method. To estimate the error variance of X using Eq. ~~11a, we omit~~ (6a), we calculate examples of the error terms ~~COV_{err}(X, Z), M(True, X_{err}), and M(True, X_{err} + Z_{err})~~for $a = 0$ and $a = 0.5$. The COV_{err}(X, Z) term was already shown in Fig. 2, and is the same in the 2CH method as in the 3CH method. Figures 5 and 6 show profiles of the other two terms for an error correlation between X and Z of $a = 0$ and 0.5. In our error model X_{err} and Z_{err} are uncorrelated with True, so the non-zero values in Figs. 5–6 are a result of the finite data set (1460 in this example). The COV_{err}(X, Z) (Fig. 2) increases as a increases, reaching a maximum value in the upper troposphere of about 800 $\%^2$ for $a = 0.5$. The terms involving True and ~~he~~ the errors in X,Y and Z in Figs. 5 and 6 do not change in magnitude

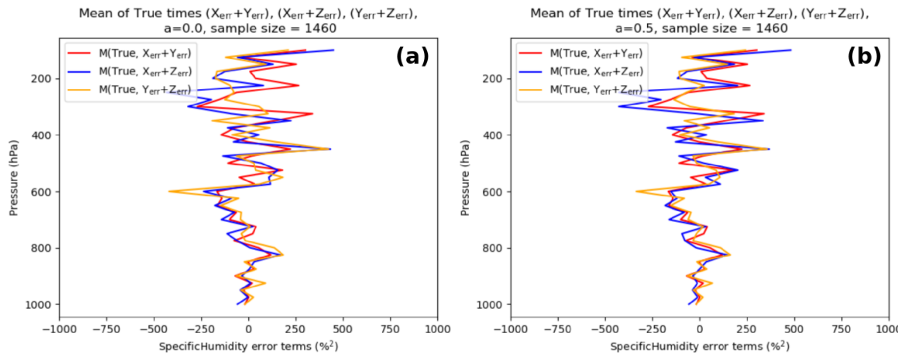

Figure 5. Terms involving products of True and ($\mathrm{X_{err}}$ + $\mathrm{Z_{err}}$), ($\mathrm{X_{err}}$ + $\mathrm{Y_{err}}$ Yerr), and ($\mathrm{Y_{err}}$ + $\mathrm{Z_{err}}$) for (a) $a = 0$ and (b) $a = 0.5$. Note that the magnitudes of True and ($\mathrm{X_{err}}$ + $\mathrm{Y_{err}}$) and True and ($\mathrm{Y_{err}}$ + $\mathrm{Z_{err}}$) do not depend on the correlation between $\mathrm{X_{err}}$ and $\mathrm{Z_{err}}$. All error terms are normalized.

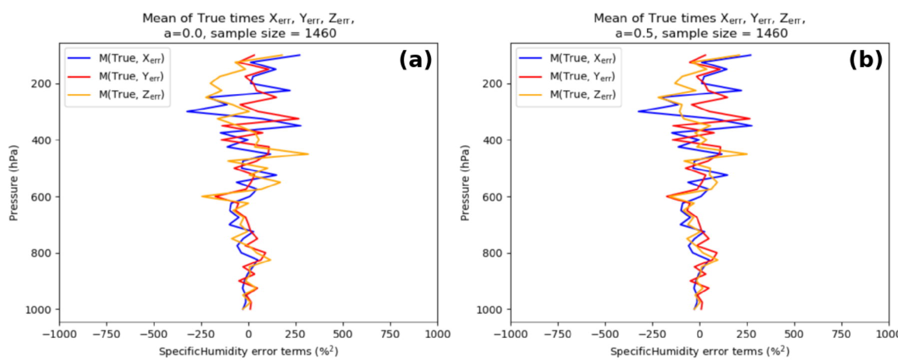

Figure 6. Mean of products of True and $\mathrm{X_{err}}$, $\mathrm{Y_{err}}$, and $\mathrm{Z_{err}}$ for (a) $a = 0$ and (b) $a = 0.5$. All error terms are normalized.

with an increasing value of a, but they increase in amplitude with height and are significantly large ($\sim$300 %^{2}) compared to the $\mathrm{COV_{err}}$(X, Z). These results indicate that a large sample size is especially important in the 2CH method; even with a sample size of 1460 the random errors caused by the neglect of the covariance terms involving True and the errors in X and Z can be significant.

Figure 7 shows the exact and estimated error STD of X, Y, and Z for various combinations of the data sets and values of a. The estimated error STD for X, Y, and Z vary around the exact solutions (black lines) for all values of a if data sets with uncorrelated errors are combined: (a) $\mathrm{STD_{err}}$(X) computed with Y, (b) $\mathrm{STD_{err}}$(Y) computed with X, and (c) $\mathrm{STD_{err}}$(Z) computed with Y, and (e) $\mathrm{STD_{err}}$(Y) using Z. Note in (c) the exact $\mathrm{STD_{err}}$(Z) ~~gets smaller~~ decreases with a increasing from 0 to 1, as described previously. Figs. 7(d) and (f) show how correlated errors between the data sets affect the estimated error variances. Both the estimated $\mathrm{STD_{err}}$(X) and $\mathrm{STD_{err}}$(Z) become too small when the data sets X and Z are combined and the value of a is increased. The exact solutions for all values of a in (f) are in black.

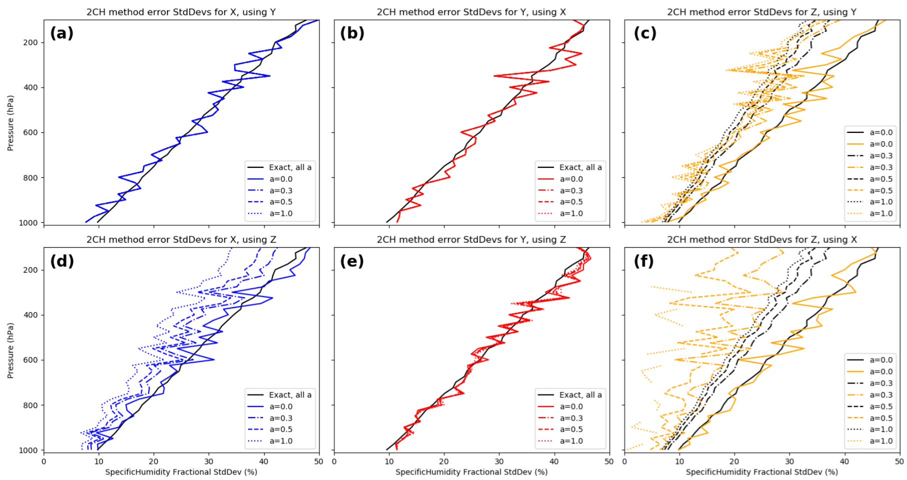

Figure 7. Two 2CH results each for the normalized error STD for X (a and d), Y (b and e) and Z (c and f), depending on the second data set. Profiles corresponding to several values of correlation parameter a are included. The solid black profile is the exact error STD profile for X and Y for all values of a, and for Z when $a = 0$.

4.4 Comparison of 3CH and 2CH methods using the error model

4.4.1 Effect of random errors in the error model

Figure 8 compares the normalized error estimates from the 3CH and 2CH methods for $a = 0.5$. In Fig. 8a (3CH) the solid lines denote exact error STD of X, Y and Z. The exact error STD are the same for X and Y, and are smaller for Z as discussed earlier. The estimated error STD (dashed lines) are less than the exact values for X and Z and greater than the exact values for Y, also as discussed earlier. In Fig. 8b (2CH), solid lines denote a combination of data sets with uncorrelated errors (where error terms are neglected, but are non-zero due to a finite sample size), while dashed lines indicate combinations of data sets with correlated errors and neglected error terms. We see that for the 3CH method and 2CH method ~~,~~ all estimates are similar~~(~~, with the exception of $\mathrm{STD_{err}(Y)}$, which is affected in the 3CH method by correlation of $\mathrm{X_{err}}$ and $\mathrm{Z_{err}}$~~)~~. However, the profiles of the 2CH estimates show considerably more noise than those of the 3CH estimates, which is a consequence of the larger magnitude of the neglected error terms in the 2CH method (see Sects. 2.3 and 4 above).

We next consider the effect of the sample size on the error estimates from the 3CH and 2CH methods. We repeat the calculations from both methods by using a subset of the 1460 samples used in the above calculations. We created the subset by selecting every tenth sample from the complete set, giving a sample size of 146.

Figure 9 shows the same estimates as those in Fig. 8, except for the much smaller sample size of 146. For the smaller sample size, the noise increases for both methods. However, the effect on the 3CH method is less than on the 2CH method, and the differences in the estimates are still clearly visible in the 3CH method. For some of our comparison data sets using real data

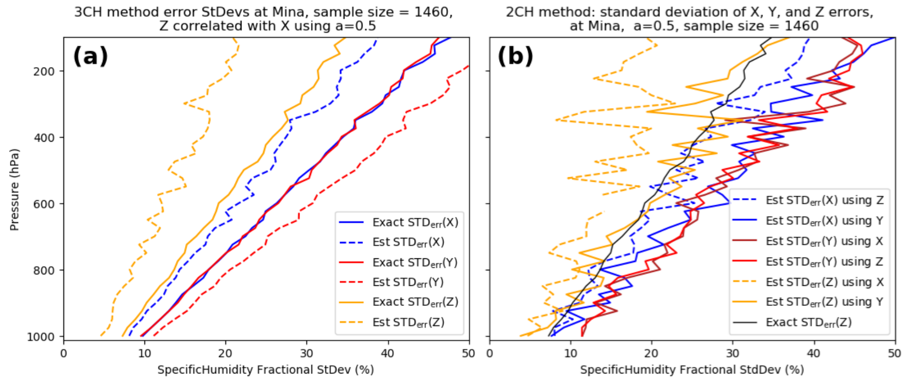

Figure 8. Estimated and exact normalized error standard deviations for 3CH method (a) and 2CH method (b) for a correlation of Z and X errors of 0.45 ($a = 0.5$). (a) Exact STD errors for X, Y and Z given by blue, red and orange solid lines respectively. Estimated error STD error given by dashed lines of same color. (b) Exact error STD of Z given by solid black line. Estimated error STD of X using Z and Y given by blue lines, estimated error STD of Y using X and Z given by red lines, and estimated error STD of Z using X and Y given by orange lines. For all colored lines, error terms are neglected. Solid lines indicate estimates from combinations of data sets with uncorrelated errors, dashed lines indicate estimates from combinations of data sets with correlated errors and hence larger error terms.

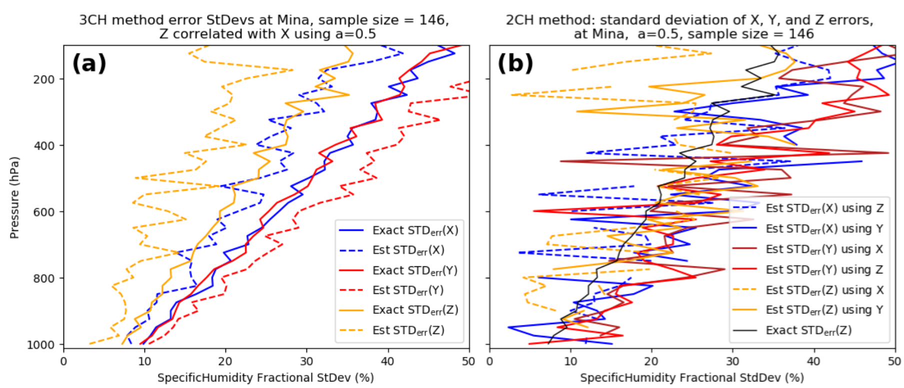

Figure 9. Same as Fig. 8 except for a smaller sample size of 146. All profiles become much noisier, with the 2CH method showing significantly more noise than the 3CH method.

(Anthes and Rieckh, 2018), the sample size n is less than 100 in the lower and upper troposphere; hence the noise due to small sampling size is likely to be significant in the estimates for these regions.

4.4.2 Effect of bias errors in the error model

To investigate the effect of both random and systematic errors, we add a known bias ϵ to data set Z.

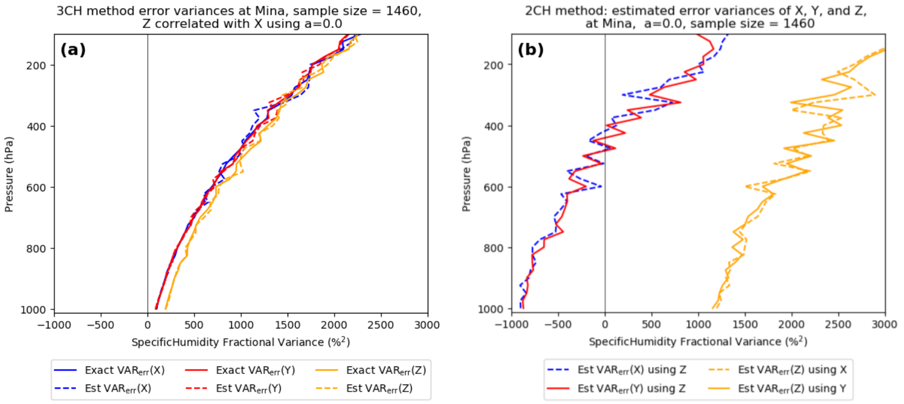

Figure 10. Effect of adding a constant bias of 10 % to Z for the 3CH (a) and 2CH (b) methods, for no correlation between X and Z errors (i.e. $a = 0$). The 3CH method yields reasonable estimates for all variables, while the estimated error variances from the 2CH method are erroneously large for Z (orange) and erroneously low X and Y (blue and red).

For the 3CH method, Eq. (3a) becomes

$$\mathrm{VAR_{err}(X)_{est}} = {}^{1}/_{2}[\mathrm{MS(X-Y) + MS(X-(Z{+}\epsilon)) - MS(Y-(Z{+}\epsilon))}] \tag{17}$$

which may be expanded to

$$\mathrm{VAR_{err}(X)_{est}} = {}^{1}/_{2}[\mathrm{MS(X-Y) + MS(X-Z) - MS(Y-Z)}] + \mathrm{M}[\epsilon(\mathrm{Y_{err} - X_{err}})] \tag{18}$$

which is equal to Eq. (3a) plus the bias term. For random errors and a large sample size the means of $\mathrm{X_{err}}$ and $\mathrm{Y_{err}}$ will be very small and the difference of the mean errors will also be small. Thus the bias term will tend toward zero as the sample size increases and the effect of a bias error in Z is minimal in the 3CH method.

This is not the case in the 2CH method, as shown below. Equation (6a) becomes

$$\mathrm{VAR_{err}(X)_{est}} = \mathrm{MS(X)} - {}^{1}/_{4}\,[\,\mathrm{MS(X + (Z{+}\epsilon)) - MS(X - (Z{+}\epsilon))}] \tag{19}$$

which can be expanded to

$$\mathrm{VAR_{err}(X)_{est}} = \mathrm{MS(X)} - {}^{1}/_{4}\,[\mathrm{MS(X+Z) - MS(X-Z)}] - \mathrm{M}(\epsilon \mathrm{X}) \tag{20}$$

which is equal to Eq. (6a) plus a bias term.

As noted above, the bias term in the 3CH method can be close to zero and therefore not affect the computed $\mathrm{VAR_{err}(X)}$ significantly. The bias term of the 2CH method will stabilize with higher sample sizes around the product of the bias (ϵ) and the M(X) and will significantly influence the estimated error variance. A positive bias in the data set Z will cause a negative error in $\mathrm{VAR_{err}(X)}$, and a negative bias in the data set Z will cause a positive error in $\mathrm{VAR_{err}(X)}$.

For the simulated data set, Eq. (12) becomes

$$\mathrm{Z_{err}} = \frac{a \cdot \mathrm{X_{err}} + \mathrm{Q_{err}}}{1+a} + \epsilon, \tag{21}$$

while $\mathrm{X_{err}}$ and $\mathrm{Y_{err}}$ stay the same.

Figure 10 shows the effect of adding a constant bias of 10 % to Z for the 3CH and 2CH methods for no correlation of random errors ($a = 0$). For the 3CH method (Figure 10a), results still look reasonable, and the estimated and exact solutions overlap well. For the 2CH method (Figure 10b), a bias of 10 % strongly impacts the estimated error variances. For Z, they are much too high, and for X and Y they are much too low. Recalling that the variables X, Y, Z and ϵ are all normalized by the sample mean $\bar{q}$, M(X) in Eq. (20) is approximately 100 %, so the term ϵ M(X) is approximately 100 ϵ %^{2}. For $\epsilon = 10\,\%$, this term is 1000 %^{2}, which is the offset we see for X, Y, and Z in Fig. 10b.

5 Estimates of error variances using 3CH and 2CH methods and real observations

Anthes and Rieckh (2018) showed a large number of error variance estimates using real data and the 3CH method at four radiosonde (RS) locations in the Pacific Ocean region. Here we show a few examples of how the 2CH method compares with these 3CH estimates.

We use co-located data of radio occultation (RO), RS, NCEP Global Forecast System (GFS), and ERA-Interim (ERA) at Minamidaitojima (Mina), which is located on Okinawa at 25.6°N 131.5°W, and is one of the four RS stations studied by Anthes and Rieckh (2018) and Rieckh et al. (2018). We use the RO-Direct method for computing RO specific humidity (uses GFS NCEP Global Forecast System temperature to compute specific humidity q from observed RO refractivity). Details about the co-location criteria and specific humidity retrieval are described by Rieckh et al. (2018).

To compute error variances using the 2CH method, Eq. (6a) is applied for various combinations of the four data sets. Data pairs at each level are only used if data are available for all four data sets. All data sets are interpolated to a common 25 hPa grid. In the 2CH method, Eq. 11a is applied for various combinations of the four data sets. Figure 11a shows the number of co-located profiles per pressure level. Figure 11b shows the normalized RS specific humidity values for these profiles.

(a) Number of co-located measurements (data pairs) per pressure level for RO, RS, GFS and ERA. (b) Normalized q values for radiosondes (RS) at Mina, 2007. The normalized q values are computed as $100(q - q_{\mathrm{ERA}})$/CLIMO where CLIMO is the annual mean value of q for 2007. Profiles cut off at 250hPa because RS data are not reported at higher levels. At the bottom, co-located profiles thin out since RO penetration depth varies and only very few profiles are available at the lowest levels. The two model and the RO data sets are representative of similar horizontal scales (~100 km), while the radiosonde data are in-situ point measurements and therefore represent a much smaller horizontal scale. However, many studies (e.g. Ho et al., 2010a, b; Kuo et al., 2004; Chen et al., 2011) have used radiosonde data as correlative data for verifying models, RO, and other data sets without applying corrections for representativeness errors. These results indicate that the different representative scales are not a significant source of error in the comparisons (unlike spatial and temporal sampling

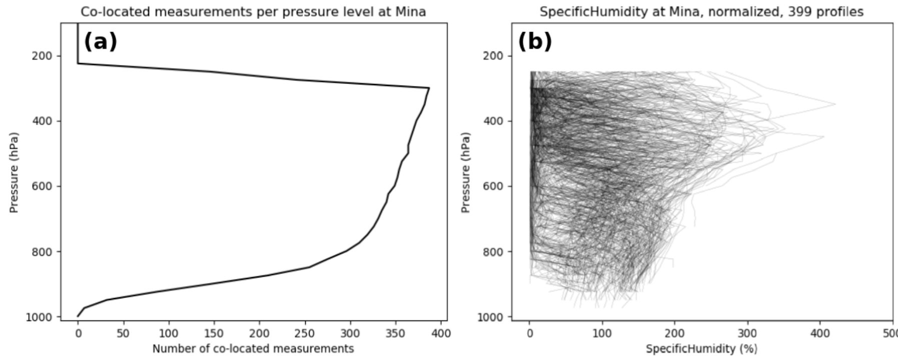

Figure 11. (a) Number of co-located measurements (data pairs) per pressure level for RO, RS, and ERA. (b) Normalized q values for radiosondes (RS) at Mina, 2007. The normalized q values are computed as $100 \cdot q/\bar{q}$ where $\bar{q}$ is the annual mean value of q for 2007. Profiles cut off at 250 hPa because RS data are not reported at higher levels. At the bottom, co-located profiles thin out since RO penetration depth varies and only very few profiles are available at the lowest levels.

errors resulting from the time and spatial differences between the data sets, which we correct for). However, any representativeness errors are included in the error estimates using either the 2CH or 3CH method.

All four data sets have some degree of unknown bias for certain locations, altitudes, or atmospheric conditions; none of them represent the ultimate "truth" and there is no standard atmospheric data set for calibration. However, they have all been compared to other models or observations to one degree or another. We investigated the effect of biases in the related paper Anthes and Rieckh (2018) by comparing the 3CH method to the triple co-location (TC) method (Stoffelen, 1998; Vogelzang et al., 2011) . The results using the TC method, in which the data sets were calibrated using the ERA-Interim data set as a calibration reference, were very similar to those using the 3CH method.

We use our results from the 3CH method for real data (Anthes and Rieckh, 2018) to evaluate the results of the 2CH method. Figure 12a shows the 3CH estimated error variances for specific humidity for ERA, ~~GFS,~~ RS and RO using three independent equations (Anthes and Rieckh, 2018). The mean value of the estimates is given by the solid line and the STD of the estimates about the mean by the shading. The 3CH estimates are considered reasonably accurate for the reasons given in Anthes and Rieckh (2018), namely that the magnitude and shape of the estimates for refractivity agree with other independent refractivity error estimates (so we assume that the method works just as well for humidity as for refractivity), and that the results are consistent for the four different RS stations studied.

The other panels in Fig. 12 show the 2CH estimates of the error variance for ERA (b), RO (c) and RS (d) using various pairs of observations. We consider the 2CH estimates unrealistic for these data sets. The magnitudes reach values that are negative or up to five times the magnitudes of the 3CH method, which is considered unrealistically large. The profiles of the estimates of error variances of ERA, RS and RO are also quite different depending on the pairs of observations used, unlike the 3CH method is which all combinations of observations give similar profiles. Similar results were obtained for 2CH estimates of refractivity (not shown).

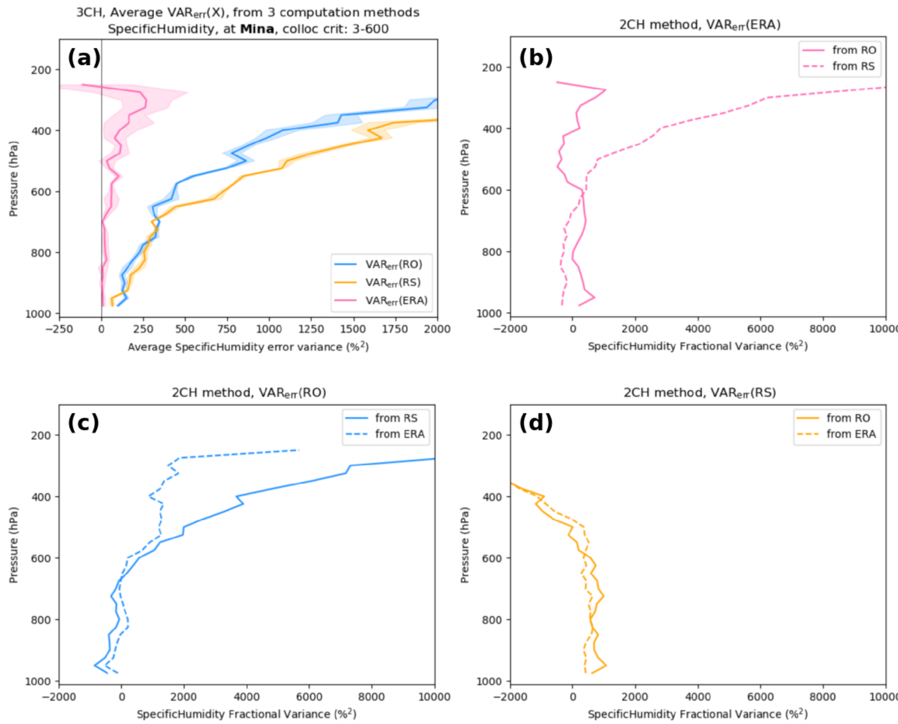

Figure 12. Estimated error variances for specific humidity: (a) results from the 3CH method for ERA (purple), ~~GFS (gray),~~ RO (blue), and RS (orange). Other three panels: 2CH method estimates for ERA (b), RO (c) and RS (d).

Thus we find that the 2CH method produces estimates of the error variances for specific humidity that are quite different from those of the 3CH method. ~~This is somewhat surprising based on the comparison of the two methods using the random error model, in which the results were similar except for the greater noise in the 2CH method, a result of the larger neglected error terms in the 2CH method.~~ We suspected that the cause for the different behavior using real data might lie in the different treatment of bias errors in the ~~real data in the~~ 2CH and 3CH method. To investigate this hypothesis, we considered the effect of ~~adding a bias error ϵ to Z.~~ an empirically based bias in the simulated data.

6 ~~Comparison of 3CH and 2CH methods including a bias~~

5.1 ~~Effect of a~~ Simulating the observed real data bias in the ~~error model~~2CH method

~~In order to investigate the effect of a bias in one of our data sets, we go back to the derivations of error variance terms~~ As shown in Sect. ~~2. For the 3CH method, Eq. ?? becomes~~

$$\mathrm{VAR_{err}(X)_{est}} = {}^1/_2[\mathrm{MS(X-Y) + MS(X-(Z+\epsilon)) - MS(Y-(Z+\epsilon))}]$$

which may be expanded to

$$\mathrm{VAR_{err}(X)_{est} = M(X^2) - M(X^2 - XY - XZ + YZ + \epsilon[M(Y_{err}) - M(X_{err})]}.$$

For random errors and a large sample size the means of X_{err} and Y_{err} will be very small and the difference of the mean errors will also be small . Thus the bias term will tend toward zero as the sample size increases and the effect of a bias error in Z is minimal in the 3CH method. However, this is not the case in the 2CH method, as shown below. Equation 11a becomes

$$\mathrm{VAR_{err}(X)_{est} = MS(X) - [\ MS(X + (Z+\epsilon)) - MS(X - (Z+\epsilon))]/4}$$

which can be expanded to

$$\mathrm{VAR_{err}(X)_{est} = MS(X) - M(XZ) - \epsilon M(X).}$$

Recalling that the variables X, Y, Z and ϵ are all normalized by CLIMO, M(X) is approximately 1 (100) or the term ϵ M(X) is approximately 100 ϵ^2.

$$\mathrm{VAR_{err}(X) = MS(X) - MS(XZ) - 100\epsilon(\%^2)} \tag{22a}$$

where ϵ is expressed as a percent.

For a bias error ϵ of 10, the error term is 1000^2, which is large compared to the true $\mathrm{VAR_{err}}$(X). For the 2CH method a positive bias in the data set Z will cause a negative error in the computed error variance of X, and a negative bias in the data set Z will cause a positive error in the computed error variance of X.

Effect of adding a constant bias of 10to Z for no correlation between X and Z errors (a) and $a = 0.5$ (b). For $a = 0$, the estimated error variance of Z is increased (orange), and the estimated error variances for X and Y computed with Z (blue dashed and red solid) are erroneously low. Adding correlation of X and Z errors to the bias in (b) creates an even more complicated picture.

Figure 10 shows the effect of adding a constant bias of 10to Z for no correlation of random errors ($a = 0$) and positive correlation between random errors of X and Z given by $a = 0.5$. The correct error variance profiles are given by the solid blue ($\mathrm{VAR_{err}}$(X) using Y) and dashed red profiles ($\mathrm{VAR_{err}}$(Y) using X. For no correlation of random errors (Fig. 10a), the effect of adding a constant bias of 10to Z is to produce estimated error variances of X using Z and Y using Z that are much too low. Conversely, the estimates of error variance of Z using X or Y are much too high. When the random errors of X and Z are correlated (Fig. 10b), similar bias 4.4.2, even small bias errors in one of the data sets can cause large errors in the estimated profiles involving Z are evident, but the correlation produces more noise in the upper troposphere.

5.2 Simulating the observed real data bias in the 2CH method

2CH method. To see if a bias in our real data could explain the very different estimates of the error variances shown in Figs. 12b–12d for the 2CH method, we set up empirically based bias profiles in the simulated data. These match approximately the observed differences of RS and RO from ERA as found by Rieckh et al. (2018) in the real data sets (supplement, Figure S5

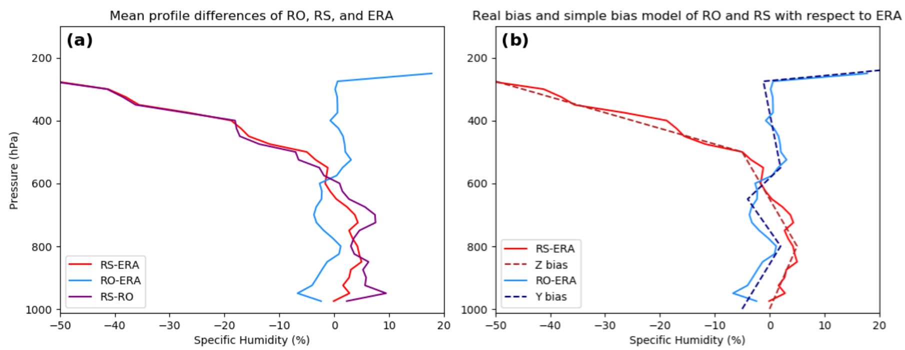

Figure 13. (a) Annual mean normalized profiles of RS – ERA, RO – ERA, and RS – RO. (~~solid red~~b) same for RS – ERA and RO – ERA only (solid ~~blue~~lines), and their respective empirical bias profiles (dashed lines) based on the mean profiles.

panel 4 for RS; Figure S6 panel 1 for RO). For the tests we consider ERA the truth. We computed the specific humidity annual mean biases of RO and RS for each level as 100[Mean(RO) – ~~CLIMO~~Mean(ERA)]/~~CLIMO~~ Mean(ERA) and 100[Mean(RS) – ~~CLIMO~~Mean(ERA)]/~~CLIMO~~Mean(ERA), respectively. We used these results to create a simple ~~mean~~ additive bias for both Y and Z. The biases are depicted in Fig. 13 (dashed lines), along with the real RS and RO annual mean biases (solid lines). More specifically, the bias used for Y (simulating the RO bias) varies linearly between pressure levels as:

–5 to 2 % from 1000 to 800 hPa

2 to –4 % from 800 to 650 hPa

–4 to 2 % from 650 to 550 hPa

2 to ~~0~~ –2 % from 550 to 275 hPa

–2 to 18 % from 275 to 250 hPa

The bias used for Z (simulating the RS bias) varies linearly between pressure levels as:

0 to 5 % from 1000 to 800 hPa

5 to –5 % from 800 to 500 hPa

–5 to -55 % from 500 to 250 hPa

The respective bias is added to both Y and Z when the data are created from the True profiles:

X = True + X_{err}; corresponds to ERA

Y = True + Y_{err}; corresponds to RO

Z = True + Z_{err}; corresponds to RS

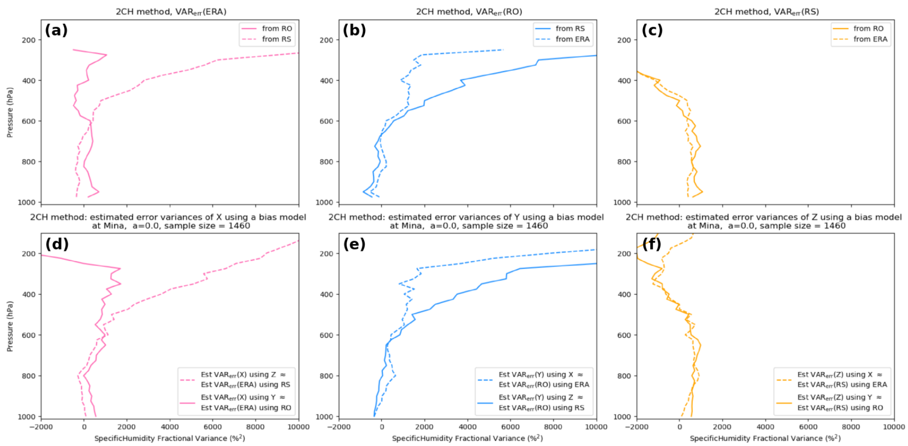

Figure 14. 2CH method error variances for ERA, RO, and RS (top row) and for simulated data using the specified empirical bias profiles (bottom). Correlation between X and Z errors is zero for this experiment.

We use these specified biased data sets to compute error variances via the 2CH method. Results are shown in Fig. 14, with the results of the real data in the top row (a–c), and the results of the simulated bias data in the bottom row (d–f).

The error variances ~~for X (Fig. 14d) look similar to its corresponding real data set ERA (Fig. 14a). The error variance estimates for X using Y (solid) versus Z (dashed) agree in their overall shape to ERA using RO (solid) versus RS (dashed). Since our empirical bias model is very simple, agreement between X and ERA estimates is not perfect. Similarities between the real (top) and simulated (bottom) estimates can also be seen for RO and Y (Figs. 14b and 14e). Differences between the real and simulated data results are largest for RS and Z (Figs. 14c and 14f), especially when RS and RO (Z and Y) are combined (solid lines)~~of the simulated biased data sets (bottom row) are quite similar to the error variances estimated with the real data (top row), indicating that biases are largely responsible for the results using real data and how they vary from the 3CH method results.

6 Summary and conclusions

In this study we compared two methods for estimating the error variances of multiple data sets, the three-cornered hat (3CH) and two-cornered hat (2CH) methods. Using a specified error model in which we could vary the degree of correlation between two data sets as well as specifying bias errors, we examined the sensitivity of the 3CH and 2CH methods to random and bias errors. For the error model, we added known random or bias errors to 1460 specific humidity profiles (considered truth) obtained from the ERA-Interim reanalysis (ERA) over a subtropical radiosonde station in Japan. We compared the effect of neglecting various error covariance and other error terms in the 3CH and 2CH methods on the estimated error variances and standard deviations. We also considered the effect of a finite sample size on the estimates by repeating the calculations using a subset of 146 of the total 1460 profiles. We found that the 3CH method was less sensitive to the neglected error terms for various random error correlations than the 2CH method. We also showed that the effect of bias errors on one of the data sets had a relatively small effect on the 3CH method, but produced much larger errors in the 2CH method.

We also compared the 3CH and 2CH methods using real radiosonde (RS)~~and~~, radio occultation (RO)~~data, as well as GFS model~~, and ERA data. We find that the 3CH method produces more consistent and accurate results than the 2CH method when using real data. The 2CH method produced very different estimates of the error variance of ERA depending on which observational data set (e.g. RS or RO) was used in the comparison. Using an empirical bias model based on observed RS and RO difference from ERA during 2007, we showed that these differences in error variance estimates were likely caused by different biases in the RS and RO data. The effect of bias errors is shown to give unrealistic results using the 2CH method.

Code availability. Code will be made available by the author upon request.

Data availability. Data can be made available from authors upon request.

Appendix A: Derivation of the 2CH method

In the 2CH method, sums and differences of two data sets X and Z are used to derive the error variances of X and Z.

$$\mathrm{X} = \mathrm{True} + \mathrm{X_{err}} \tag{A1}$$

$$\mathrm{Z} = \mathrm{True} + \mathrm{Z_{err}} \tag{A2}$$

$$(\mathrm{X}+\mathrm{Z})^2 = 4\mathrm{True}^2 + 4\mathrm{True}\,(\mathrm{X_{err}}+\mathrm{Z_{err}}) + (\mathrm{X_{err}}+\mathrm{Z_{err}})^2 \tag{A3}$$

(A4)

Equation (A3) is summed over all the data pairs to get

$$\frac{1}{n}\sum(\mathrm{X}+\mathrm{Z})^2 = 4\frac{1}{n}\sum\mathrm{True}^2 + \frac{1}{n}\sum\left[(\mathrm{X_{err}}+\mathrm{Z_{err}})^2 + 4\mathrm{True}\,(\mathrm{X_{err}}+\mathrm{Z_{err}})\right] \tag{A5}$$

Defining the following expressions:

$\mathrm{M(X)} = {}^{1}/_{n}\sum \mathrm{X}$

$\mathrm{M(X,Y)} = \mathrm{M(X \cdot Y)} = {}^{1}/_{n}\sum \mathrm{XY}$

Eq. (A5) becomes

$$\begin{aligned}\mathrm{MS(X+Z)} =& 4\mathrm{MS(True)} + \mathrm{VAR_{err}(X)} + \mathrm{VAR_{err}(Z)} \\ &+ 2\mathrm{COV_{err}(X,Z)} + 4\mathrm{M(True, X_{err}+Z_{err})}\end{aligned} \tag{A6}$$

We then subtract Z from X and square the difference to get

$$MS(X-Z) = VAR_{err}(X) + VAR_{err}(Z) - 2COV_{err}(X, Z) \quad (A7)$$

Finally, we solve for MS(True) by subtracting Eq. (A7) from Eq. (A6)

$$4MS(True) = MS(X+Z) - MS(X-Z) - 4COV_{err}(X, Z) - 4M(True, X_{err} + Z_{err}) \quad (A8)$$

By squaring the expression $X = (True + X_{err})$ we get the exact expression for the $VAR_{err}(X)$

$$VAR_{err}(X) = MS(X) - MS(True) - 2M(True, X_{err}) \quad (A9)$$

and substituting for MS(True) from Eq. (A8) gives

$$\begin{aligned} VAR_{err}(X) = & MS(X) - {}^{1}/_{4}[MS(X+Z) - MS(X-Z)] \\ & - 2M(True, X_{err}) + COV_{err}(X, Z) + M(True, X_{err} + Z_{err}) \end{aligned} \quad (A10)$$

Similarly, we obtain

$$\begin{aligned} VAR_{err}(Z) = & MS(Z) - {}^{1}/_{4}[MS(X+Z) - MS(X-Z)] \\ & - 2M(True, Z_{err}) + COV_{err}(X, Z) + M(True, X_{err} + Z_{err}) \end{aligned} \quad (A11)$$

Author contributions. Both authors contributed equally to the ideas and conceptual development. The first author computed the results.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors were supported by NSF-NASA grant AGS-1522830. We thank Shay Gilpin for Fig. 4 and the two reviewers for their comments on the original draft. We thank Eric DeWeaver (NSF) and Jack Kaye (NASA) for their long-term support of COSMIC.

References

Anthes, R. and Rieckh, T.: Estimating observation and model error variances using multiple data sets, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-487, in review, 2018.

Braun, J., Rocken, C., and Ware, R.: Validation of line-of-sight water vapor measurements with GPS, Radio Sci., 36, 459–472, 2001.

Chen, S.-Y., Huang, C.-Y., Kuo, Y.-H., and Sokolovskiy, S.: Observational Error Estimation of FORMOSAT-3/COSMIC GPS Radio Occultation Data, Mon. Wea. Rev., 139, 853–865, https://doi.org/10.1175/2010MWR3260.1, 2011.

Collard, A. D. and Healy, S. B.: The combined impact of future space-based atmospheric sounding instruments on numerical weather-prediction analysis fields: A simulation study, Quart. J. Roy. Meteor. Soc., 129, 2741–2760, https://doi.org/10.1256/qj.02.124, 2003.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quart. J. Roy. Meteor. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Desroziers, G. and Ivanov, S.: Diagnosing and adaptive tuning of observation-error parameters in a variational assimilation, Quart. J. Roy. Meteor. Soc., 127, 1433–1452, https://doi.org/10.1002/qj.49712757417, 2001.

Ekstrom, C. R. and Koppang, P. A.: Error Bars for Three-Cornered Hats, IEEE Trans. Ultrason. Ferroelect. Freq. Contr., 53, 876–879, https://doi.org/10.1109/TUFFC.2006.1632679, 2006.

Gray, J. E. and Allan, D. W.: A method for estimating the frequency stability of an individual oscillator, Atlantic City, New Jersey, May 29–31, 1974.

Griggs, E., Kursinski, E., and Akos, D.: An investigation of GNSS atomic clock behaviour at short time intervals, GPS Solut., 18, 443–452, https://doi.org/10.1007/s10291-013-0343-7, 2014.

Griggs, E., Kursinski, E., and Akos, D.: Short-term GNSS satellite clock stability, Radio Sci., 50, 813–826, https://doi.org/10.1002/2015RS005667, 2015.

Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., and Wagner, W.: Recent advances in (soil moisture) triple collocation analysis, Int. J. Appl. Earth Obs. and Geoinf., 45, 200–211, https://doi.org/10.1016/j.jag.2015.09.002, 2016.

Ho, S.-P., Kuo, Y.-H., Schreiner, W., and Zhou, X.: Using SI-traceable Global Positioning System radio occultation measurements for climate monitoring, Bull. Amer. Meteor. Soc., 91, S36–S37, 2010a.

Ho, S.-P., Zhou, X., Kuo, Y.-H., Hunt, D., and Wang, J.-H.: Global evaluation of radiosonde water vapor systematic biases using GPS radio occultation from COSMIC and ECMWF analysis, Remote Sensing, 2, 1320–1330, https://doi.org/10.3390/RS2051320, 2010b.

Kuo, Y.-H., Wee, T.-K., Sokolovskiy, S., Rocken, C., Schreiner, W., Hunt, D., and Anthes, R. A.: Inversion and error estimation of GPS radio occultation data, J. Meteor. Soc. Japan, 82, 507–531, 2004.

Kursinski, E. R., Hajj, G. A., Schofield, J. T., Linfield, R. P., and Hardy, K. R.: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System, J. Geophys. Res., 102, 23 429–23 465, https://doi.org/10.1029/97JD01569, 1997.

Luna, D., Pérez, D., Cifuentes, A., and Gómez, D.: Three-Cornered Hat Method via GPS Common-View Comparisons, IEEE Trans. Instrument. Measure., 66, 2143–2147, https://doi.org/10.1109/TIM.2017.2684918, 2017.

O'Carroll, A. G., Eyre, J. R., and Saunders, R. S.: Three-way error analysis between AATSR, AMSR-E, and in situ sea surface temperature observations, J. Atmos. Oceanic Tech., 25, 1197–1207, https://doi.org/10.1175/2007JTECHO542.1, 2008.

Rieckh, T., Anthes, R., Randel, W., Ho, S.-P., and Foelsche, U.: Evaluating tropospheric humidity from GPS radio occultation, radiosonde, and AIRS from high-resolution time series, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-486, accepted, 2018.

Riley, W. J.: http://www.wriley.com/3-CornHat.htm, 2003.

Roebeling, R. A., Wolters, E. L. A., Meirink, J. F., and Leijnse, H.: Triple collocation of summer precipitation retrievals from SEVIRI over Europe with gridded rain gauge and weather radar data, Journal of Hydrometeorology, 13, 1552–1566, https://doi.org/10.1175/JHM-D-11-089.1, 2012.

Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, J. Geophys. Res., 103, 7755–7766, https://doi.org/10.1029/97JC03180, 1998.

Su, C.-H., Ryu, D., Crow, W. T., and Western, A. W.: Beyond triple collocation: Applications to soil moisture monitoring, J. Geophys. Res., 119, 6419–6439, https://doi.org/10.1002/2013JD021043, 2014.

Valty, P., de Viron, O., Panet, I., Camp, M. V., and Legrand, J.: Assessing the precision in loading estimates by geodetic techniques in Southern Europe, Geophys. J. Int., 194, 1441–1454, https://doi.org/https://doi.org/10.1093/gji/ggt173, 2013.

Vogelzang, J., Stoffelen, A., Verhoef, A., and Figa-Saldaña, J.: On the quality of high-resolution scatterometer winds, J. Geophys. Res., 116, https://doi.org/10.1029/2010JC006640, 2011.

von Engeln, A. and Nedoluha, G.: Retrieval of temperature and water vapor profiles from radio occultation refractivity and bending angle measurements using an Optimal Estimation approach: a simulation study, Atmos. Chem. Phys., 5, 1665–1677, https://doi.org/10.5194/acp-5-1665-2005, 2005.

Wang, B.-R., Liu, X.-Y., and Wang, J.-K.: Assessment of COSMIC radio occultation retrieval product using global radiosonde data, Atmos. Meas. Tech., 6, 1073–1083, https://doi.org/10.5194/amt-6-1073-2013, 2013.