Research Article Open Access

X. Li*

Modeling the North American vertical datum of 1988 errors in the conterminous United States

https://doi.org/10.1515/jogs-2018-0001
Received August 30, 2017; accepted December 20, 2017

Abstract: A large systematic difference (ranging from −20 cm to +130 cm) was found between NAVD 88 (North American Vertical Datum of 1988) and the pure gravimetric geoid models. This difference not only makes it very difficult to augment the local geoid model by directly using the vast NAVD 88 network with state-of-the-art technologies recently developed in geodesy, but also limits the ability of researchers to effectively demonstrate the geoid model improvements on the NAVD 88 network. Here, both conventional regression analyses based on various predefined basis functions such as polynomials, B-splines, and Legendre functions and the Latent Variable Analysis (LVA) such as the Factor Analysis (FA) are used to analyze the systematic difference. Besides giving a mathematical model, the regression results do not reveal a great deal about the physical reasons that caused the large differences in NAVD 88, which may be of interest to various researchers. Furthermore, there is still a significant amount of no-Gaussian signals left in the residuals of the conventional regression models. On the other side, the FA method not only provides a better fit of the data, but also offers possible explanations of the error sources. Without requiring extra hypothesis tests on the model coefficients, the results from FA are more efficient in terms of capturing the systematic difference. Furthermore, without using a covariance model, a novel interpolating method based on the relationship between the loading matrix and the factor scores is developed for predictive purposes. The prediction error analysis shows that about 3-7 cm precision is expected in NAVD 88 after removing the systematic difference.

Keywords: Factor Analysis; Geoid models; GPS/Leveling; North American Vertical Datum of 1988

***Corresponding Author: X. Li:** Data Solution Technology Inc., 1315 East-West Highway, Silver Spring, Maryland, U.S.A., E-mail: Xiaopeng.Li@noaa.gov
orcid.org/0000-0002-3854-8130

1 Introduction

The NAVD 88 in the Conterminous United States (CONUS) area contains over 20,000 GPS/Leveling benchmarks (at least another 10,000 points are expected to be available for the final NGS hybrid geoid model, i.e. Geoid19), making it one of the largest continental vertical datums in the world, at least in terms of size. In addition to serving as the local datum, the NAVD 88 derived geoid is often used on these benchmarks to compare to various global and local geoid models. However, in addition to the implicit datum offset, a very clear tilt with a range of 1.5 meters across the continent was found when comparing NAVD 88 with various gravimetric geoid models such as EGM2008 (Pavlis et al. 2012), EIGEN6c4 (Förste et al. 2014), and the NGS xGeoid models (xGeoid14, Roman and Li 2014; xGeoid15, Li et al. 2016; and xGeoid16) that include both the global gravity signals offered by the Gravity Recovery and Climate Experiment (GRACE; Tapley et al 2004) and the Gravity field and steady-state Ocean Circulation Explorer (GOCE; Rummel et al. 2011) and the updated local gravity field information provided by the Gravity for the Redefinition of the American Vertical Datum (GRAV-D; Smith 2007); please see Smith et al. 2013 and the NGS webpages for the technical details of the NGS experimental geoid models (https://beta.ngs.noaa.gov/GEOID/).

These large differences make it very difficult to augment the local geoid model by directly using this vast network of benchmarks with methods recently developed in geodesy, such as Prutkin and Klees (2008), Klees and Prutkin (2010), where the gravimetric geoid errors at the GPS/Leveling benchmarks can be formulated into a residual boundary value problem. It also limits the ability of researchers to effectively demonstrate the geoid model improvements at these benchmarks. Any previously reported precision of the geoid errors based on their misfits at these benchmarks are questionable and cannot be used directly as a rigorous interpretation of the geoid model precision at these NAVD 88 benchmarks.

Effectively removing the systematic errors in NAVD 88 is very necessary for many practical applications. Even after NAVD 88 is replaced by a geoid-based vertical datum

by 2022, such kind of information is still useful for understanding the errors in the vast amount of historical applications based on NAVD 88. However, the leveling defining NAVD 88 started about 100 years ago and its last nationwide adjustment was done several decades ago (Zilkoski *et al.* 1992). It is almost impossible to try to reprocess everything from the very beginning at this time. As a result, the differences are studied here directly by using numerical methods that include regression analyses (Koch 1988) and the Latent Variable Analysis (LVA) methods such as Factor Analysis (FA; Rummel 1988; Rencher 2002) in order to efficiently model these errors and try to find any possible physical explanations.

Section 2 briefly describes the differences between the gravimetric geoid models and the NAVD 88 determined geoid undulations. Several mathematical techniques based on polynomials, B-splines, and Legendre functions are used to try to model these discrepancies by the regression analyses in Section 3. In Section 4, the fundamental idea of FA is recapped first while leaving the derivation details on how to solve the problem in Appendix A for the purpose of self-completeness of the paper. Then, under the assumption that the systematic error in NAVD 88 was caused by multiple reasons, FA is employed to investigate the primary pattern of the NAVD 88 errors. In addition, a novel interpolation method is also developed and tested with real numbers. In Section 5, the results are compared with some recent gravity field models published by the International Centre for Global Earth Models (ICGEM). The comparison clearly shows the improvements of the newer models at these calibrated NAVD 88 benchmarks. Finally, the conclusions are presented in Section 6.

2 Statement of the problem

It is well known that the accuracy of joint GRACE and GOCE models is better than a few centimeters with about 100 km spatial resolution (Rummel 2012, Gruber 2014), which is approximately equivalent to degree and order 200 in the spherical harmonic domain. Hence, the gravimetric geoid models that contain these satellite models should have equivalent accuracy in the long wavelength. Furthermore, recent studies have shown that the airborne-enhanced NGS xGeoid models have about a 1-2 cm accuracy at the newly observed independent leveling lines (Smith et al 2013) and at the multi-year averaged mean altimetry passes over the Great Lakes (Li et al. 2016). However, when compared with the GPS/Leveling determined geoid values at the NAVD 88 benchmarks, all of the above mentioned geoid models show significant differences to this leveling based vertical datum. For instance, Fig. 1 shows the differences between NAVD 88 and xGeoid16B, where a clear diagonal "tilt" is identified across the continent ranging from about −20 cm in the state of Florida on the Southeast to about 130 cm in the state of Washington on the Northwest.

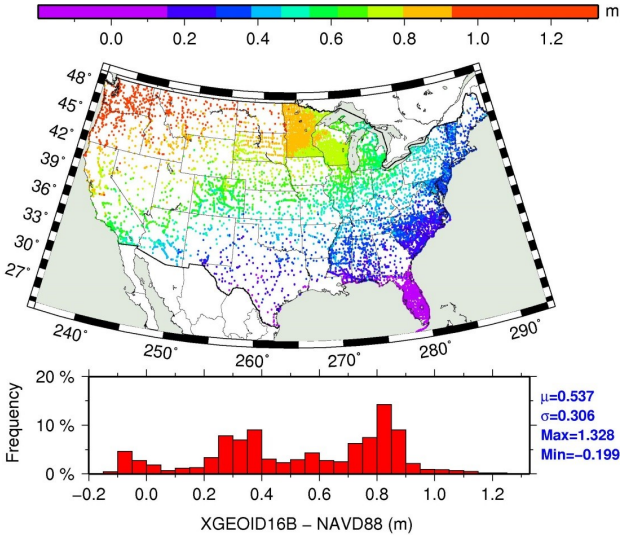

Figure 1: The geoid differences between the newest experimental geoid model, xGeoid16B, and the NAVD 88 orthometric height implied geoid heights at 21,112 GPS/Leveling benchmarks over the target area.

Both the spatial distribution and the magnitude of the differences appears to indicate that the errors must come from NAVD 88. However, unlike the Australian Height Datum that was constrained to several tide gauge stations (Featherstone and Filmer 2012), NAVD 88 is only tied to a single tidal gauge station at Rimouski, Quebec, which makes it impossible to apply the method developed by Featherstone and Filmer (2012) to resolve the current problem. If the quality of the gravimetric geoid models is believed to some extent, the "NAVD 88 systematic error" can be estimated and modeled. One easy way is just to use a satellite-only model up to certain degree, say degree and order 200, as an "error-free" reference, and model the NAVD 88 differences to this reference surface only up to this selected degree (200). Apparently, the omission errors of the reference model have to be somehow removed. Three techniques, i.e. a 200 km half wavelength Gaussian Filter, the Radial Basis Function (RBF) method (Li 2017), and the spherical harmonics are used to remove/reduce the omission effects of the reference model and try to only model the signal in the interesting band. The first two

methods can be applied directly to the NAVD 88 points. Extra gridding and global zero padding steps are required when the last technique is used. The computed NAVD 88 errors from all of these three methods are shown in Fig. 2, which shows that different methods give different answers due to the irregular data distribution, though they all show about the same trend in general. Most importantly, they all cannot tell if there is any systematic error in NAVD 88 that is above the resolution of the reference field. Moreover, we also would like to try to investigate if there are some fundamental reasons that caused the errors. Thus, a "full" gravimetric geoid model (the xGeoid16B model that includes the satellite data, airborne data, and surface data as well as the residual terrain effects) is used as a reference surface. Based on this reference, the NAVD 88 error is analyzed by using both regression analyses based on predefined basis functions and LVA methods such as FA in the following two sections.

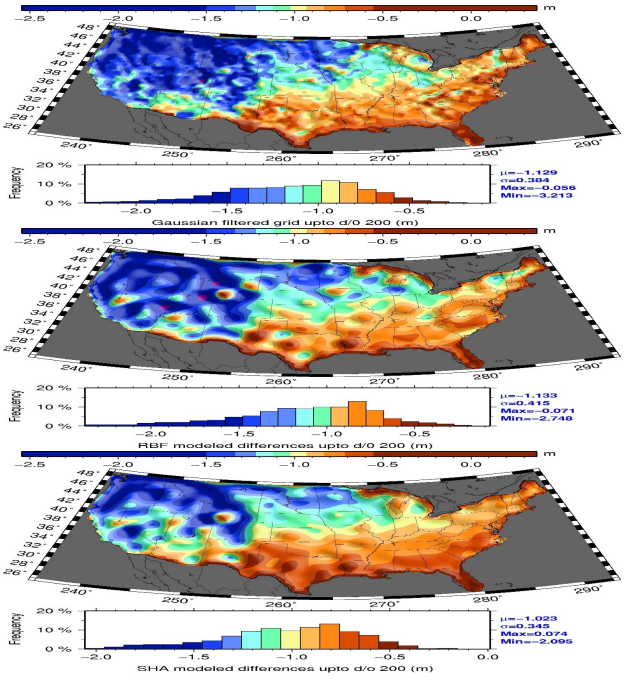

Figure 2: The modeled NAVD 88 differences with respect to GOCO05s up to degree 200 by using different approaches (upper panel Gaussian filter; middle panel RBF; lower panel SHA).

3 Regression analyses

Maybe the simplest approach is to use regression analysis to model the errors as described by Eq. (1).

$$\tilde{\zeta}_i := N_i - \left(h_i - H_i^{NAVD88}\right) = \beta_0 + \beta_1\varphi_i + \beta_2\lambda_i + \beta_3 h_i + \ldots + \epsilon_i$$
$$i = 1, \ldots, n, \quad \text{where } n = 21,112 \tag{1}$$

where the observable $\tilde{\zeta}$ is the difference between the geoid undulation N (computed from the gravimetric geoid models) and the measured one by taking the difference between the ellipsoidal height h and the orthometric height H^{NAVD88}, φ is the geodetic latitude, λ is the geodetic longitude, $\{\beta_{0,1,2\ldots k}\}$ are the coefficients needed to be estimated, and ϵ is assumed to be random noise.

The forward selection method (Hocking 1976) is used to try to find the necessary variables that will be needed to be included in Eq. (1). The analysis report of the forward selection is included in Table 1. The first column shows the models that are tested in each step. The second column gives the adjusted R square ($R^2_{adj} = 1 - \left[\frac{\sum (e_i)^2}{ss_{tot}}(n-1)\right]$ with sample size $n$and number of variables $k$as well as the sum of the model residuals $\sum (e_i)^2$ from the total sum of squares, ss_{tot}), an indicator of the model fit that increases only if the new term improves the model more than expected by chance. The standard error of the estimate ($\sqrt{\frac{\sum (e_i)^2}{n-k-1}}$) in column three is a measure of the accuracy of the model prediction. The last few columns are the t-test statistics for the significances of the coefficients of the model.

Table 1 shows clearly that the coefficients are keeping significant while the actual fit does not change too much. Changing the forward selection method into other methods such as backward selection and stepwise selection method gives similar results. This tells that the simple regression method may not be a sufficient choice for this kind of problem though a linear trend model (Roman *et al.* 2010a, 2010b, and Pavlis et al.2012) was often used to represent the NAVD 88 systematic error. Furthermore, the quantile by quantile (Q-Q) plot, in Fig. 3, shows that the residuals of this linear fitting are far away from the presumed normal distribution that has been assumed by the model described in Eq. (1). As such, some extra treatment such as the multi-matrix technique (e.g. Roman *et al.* 2004; Roman *et al.* 2010c) or the generalized Least Squares Collocation (LSC) (Klees and Prutkin 2010) have to be applied to deal with these residuals.

In addition to the polynomial functions, some other local basis functions based on B-splines and Legendre functions can be used in the regression analysis to model

Table 1: Summary of the forward selection report.

Model	Adjusted R Square	Std. Error of the Estimate	Coefficients		t	Sig.
			β_i	Std. Error		
β_0	0.837	0.124	−1.189	0.005	−223.885	$\ll 0.01$
$+\beta_1\varphi_i$			0.044	1.32E-4	329.368	$\ll 0.01$
β_0			1.789	0.012	145.881	$\ll 0.01$
$+\beta_1\varphi_i$	0.959	0.062	0.037	7.2E-5	516.016	$\ll 0.01$
$+\beta_2\lambda_i$			−.010	4.0E-5	−248.830	$\ll 0.01$
$\beta_0+\beta_1\varphi_i$			2.038	0.014	149.898	$\ll 0.01$
$+\beta_2\lambda_i$	0.961	0.060	.037	7.0E-5	534.305	$\ll 0.01$
$+\beta_3 h$			−.011	4.6E-5	−238.744	$\ll 0.01$
β_0			−4.562E-5	1.0E-6	−37.600	$\ll 0.01$
$+\beta_1\varphi_i$			0.627	0.081	7.739	$\ll 0.01$
$+\beta_2\lambda_i$	0.962	0.060	0.072	0.002	36.562	$\ll 0.01$
$+\beta_3 h$			−0.006	2.97E-4	−19.499	$\ll 0.01$
$+\beta_4\varphi_i\lambda_i$			−4.178E-5	1.0E-6	−34.139	$\ll 0.01$
			−1.27E-4	7.0E-6	−17.641	

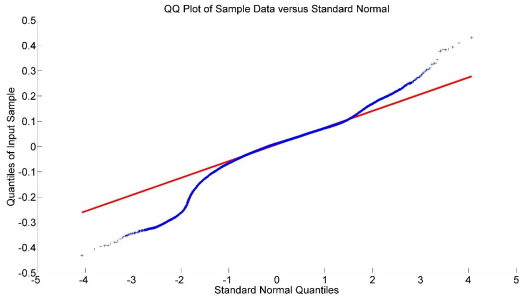

Figure 3: The Q-Q plot of the model residuals as exemplified by the blue symbols (the quantiles of the residuals versus the theoretical quantiles from a normal distribution. A straight line, as indicated by the red line, is expected if the residuals are normal.

the NAVD 88 errors; see Appendix A. However, the numerical tests show that they do not add new insight to the problem. Moreover, they all have one key drawback: they depend on predefined basis functions that lack any direct physical support for this particular application. These other methods also face under-fitting or over-fitting problems when separating the systematic component from the random errors (see Cawley and Talbot 2010 for details). Because real data are often unpredictably noisy, a coarse fitting usually produces complicated residuals. On the other hand, a high order model often has the risk of being over parameterized. Furthermore, the results from high order models are difficult to understand and properly explain.

4 Latent variable analysis (LVA)

Considering that the leveling used to define the NAVD 88 datum dates back to the early 20th century and that its last nationwide adjustment was done over 30 years ago (Zilkoski et al.1992), it would be very difficult and expensive to identify the specific reasons for the systematic errors and, consequently, rigorously formulate an effective and relatively simple mathematical model. Therefore, rather than speculating on the "true" physical background and forcing the data into some predefined basis functions, FA is used to analyze the systematic differences and to try to find some coherent physical explanation of the detected NAVD 88 errors.

4.1 A short overview of FA

FA is a process to explain observed relations among variables without knowledge of the physical causes of the changes (Cattell 1965, Rummel 1988). The rationale and mathematical developments of FA were well documented by Cattell (1965) and Rummel (1988), and tutorially described in many multivariate textbooks such as Rencher 2002. However, for the purpose of self-completeness of this paper and for the convenience of the readers, a short review is still given in the following paragraphs without heavily repeating these previous publications.

FA has two large branches. One is called Exploratory FA. The other is the Confirmatory FA. Here, we are mainly

focusing on the former and trying to find the hidden patterns in the data. The essential assumption of FA is that the observed variables are actually influenced by (fewer) factors that are not observed directly. In the literature, there are many ways to carry out this idea. For the reason of simplicity, let's assume that $\{x_{i,[1,2,\ldots,p]}\}$ are observed variables from object i, and $\{f_{i,[1,2,\ldots,q]}\}$ are the factors. Then we have the common factor model read as in Eq. (2).

$$\begin{pmatrix} x_{i,1} \\ x_{i,2} \\ \vdots \\ x_{i,p} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} + \begin{pmatrix} f_{i,1}w_{11} + f_{i,2}w_{21} + \cdots + f_{i,q}w_{q1} \\ f_{i,1}w_{12} + f_{i,2}w_{22} + \cdots + f_{i,q}w_{q2} \\ \vdots \\ f_{i,1}w_{1p} + f_{i,2}w_{2p} + \cdots + f_{i,q}w_{qp} \end{pmatrix} + \begin{pmatrix} \epsilon_{i,1} \\ \epsilon_{i,2} \\ \vdots \\ \epsilon_{i,p} \end{pmatrix} \tag{2}$$

where $\{w_{j,r};\ j = 1,\ldots,p,\ r = 1,\ldots,q\}$ are the factor loadings, $\{\epsilon_{i,j};\ i = 1,\ldots,n\}$ are the noise terms that are assumed to have zero mean and variance ψ_j (i.e. different variables have differently sized noise terms), and $E\{\epsilon_{i,j}, \epsilon_{l,m}\} = 0$ unless $i = l \& j = m$ (each observation and each variable have uncorrelated noise, which may not represent the actual cases for some real data), $\{\mu_j\}$ is the mean value of each variable, which can be easily removed from the data if we assume that the sample average is a good approximation of the population mean, and most importantly we assume the factors $\{f_{i,[1,2,\ldots,q]}\}$ are uncorrelated with a variance of 1.

Equation (2) is very similar to a linear multiple regression model. However, it has totally different meanings. This model postulates that observed measures are affected by underlying common factors ($f_{i,j}$) and unique factors ($\epsilon_{i,j}$), and that correlation patterns need to be determined (Yong and Pearce 2013). The basic idea behind this model is that FA tries to look for factors such that when these factors are extracted, there remain no inter-correlations between any pairs of $\{x_{i,[1,2,\ldots,p]}\}$, because the factors themselves will account for the inter-correlations. **Note: this does not mean that $\{x_{i,[1,2,\ldots,p]}\}$ itself are independent. It only means that all pairs of any two elements of $\{x_{i,[1,2,\ldots,p]}\}$ are conditionally independent given the value of $\{f_{i,[1,2,\ldots,q]}\}$.**

After removing the mean values by the so-called centering procedure in statistical analysis, the above system in Eq. (2) can be re-written in vector form as:

$$\vec{X}_i = \vec{F}_i^W + \vec{\epsilon}_i \tag{3}$$

with $\vec{F}_i = [f_{i,1}, f_{i,2}, \ldots, f_{i,q}]$ and $f_{i,j}$ represents subject i's score on factor j. W is a $q \times p$ matrix and it is the same for all subjects, and $\vec{X}_i$ represents the observation i that contains p variables. After stacking all n observations into a matrix, we have the following equation:

$$X = F\ W + \epsilon \tag{4}$$

FA is used to estimate the factor scores F and the loading matrix W as well as the specific errors ϵ. It is assumed that all the factor scores are uncorrelated with each other and have variance 1 and that they are uncorrelated with the noise terms. The above estimation problem can be reduced to an eigenvalue problem after noticing that n times the sample covariance matrix V has the following relationship with the loading matrix as shown in Eq. (5):

$$nV = X^T X = \epsilon^T\epsilon + \epsilon^T FW + W^T F^T \epsilon + W^T F^T FW = n\psi + nW^T W, \tag{5}$$

where ψ is a diagonal matrix whose entries are ψ_j.

The factors F are eliminated in Eq. (5), so one only needs to figure out the specific errors and the loading coefficients. There are p^2 equations, one for each entry of V and $p + pq$ unknowns in ψ and W, which means that there is no exact solution in general. There are two main methods to estimate ψ and W. The first one is called the Principal Axis Factoring (PAF). The second one is the Maximum Likelihood (ML), which needs an extra multivariate normality assumption for $\vec{X}_i$. The detailed derivations of both the PAF and ML methods are given in Appendix B.

After obtaining the estimation of the specific errors and the loading coefficients, the factor scores can be easily obtained by minimizing the mean squared error. However, unrotated factors are ambiguous (Yong and Pearce 2013). For better interpolation of the results, the loading matrix is usually rotated according to certain rules (Browne 2001). The goal of the rotation is to attain an optimal simple structure which attempts to have each variable load on as few factors as possible, but maximizes the number of high loadings on each variables (Rummel, 1988). Here the Varimax method (Kaiser 1958), a rotation that gives maximum variance of the loadings, is applied in order to concentrate the power of each individual loading coefficient.

There are two major misconceptions about the use of FA for the present application that can lead to potential misunderstandings. The first misconception is that FA requires the data to have a multi-normal (or near multinormal) distribution. However, this requirement only applies when tests of statistical significance are applied to the factor results (Rummel 1988) or a certain PDF is assumed when using the maximum likelihood method, which is recapped in Appendix B. "FA can be meaningfully

applied even to nominally scaled bi-variant (yes-no) data, the lowest and least demanding rung on the measurement ladder" (Rummel 1988). From a practical point of view, however, FA does not have any restrictions on the content of the data. The other misconception is that FA may be misunderstood to be equivalent to the Principal Component Analysis (PCA), confusion probably caused by their heavily overlapped terminologies and their algorithms which resemble one another. Indeed, they were not very distinguishable in the earlier literatures, but there are fundamental differences starting with their underlying models and goals. The components of PCA are calculated as linear combinations of the original variables, $\{x_{i,[1,2,\ldots,p]}\}$. PCA is not concerned with the error in the data; it will try to reproduce "***true-value-plus-noise***" from a small number of components (Shalizi 2009). In FA, the original variables are defined as linear combinations of the factors, $\{f_{i,[1,2,\ldots,q]}\}$. FA analyzes only the shared variances, U; error (ψ) is estimated apart from the shared variances (see Appendix B). The bottom line is that PCA is mainly designed for data reduction whereas FA is used for explanatory studies. In-depth discussions about the differences between FA and PCA were described in Rummel 1988, Cattell 1965, and more recently by Suhr 2005.

4.2 The application of FA to the NAVD 88 problem

After clearing all these hurdles, the application of FA to our problem is straightforward. The first step is to standardize or normalize the data by removing the means and dividing them by the standard deviations, which is a commonly used statistical procedure (Mulaik 1972). The standardized observations at each benchmark are put into the vector $\vec{X}_i$. The observation data includes latitude (φ), longitude (λ), ellipsoidal height (h), and the geoid difference (δN), which leads to Eq. (6):

$$\vec{X}_i = [\varphi, \lambda, h, \delta N]_i \tag{6}$$

which implies that $p = 4$. The correlation matrix V is given in Table 2. The magnitudes of correlation coefficients between the geoid error (δN) and other variables are all above 0.3. Though the correlation between the geoid error and the latitude is bigger than 0.9 (a simple rule of thumb for possible collinearity), the determinant of the correlation matrix is still much larger than the normally used threshold for collinearity, i.e. 0.00001. Furthermore, the *Bartlett's test of sphericity* (Snedecor and Cochran, 1989) has a very small p-value ($\ll 0.05$) that confirms that the data has patterned relationships in the variables.

Table 2: Correlation Matrix (Determinant = 0.023)

Correlation	φ	λ	h	δN
φ	1.000	−0.368	0.269	0.915
λ	−0.368	1.000	−0.567	−0.660
h	0.269	−0.567	1.000	0.380
δN	0.915	−0.660	0.380	1.000

Since the sample size is much bigger than 200, the scree test (eigenvalues vs number of factors, Cattell 1965) in Fig. 4 is used to determine the number of underlying factors. However, the point of inflexion is not that clear in Fig. 4. To avoid any possible drawbacks of using these methods in determining the number of factors, the FA analysis is run for both the case of 2-factor and 3-factor scenarios as suggested by (Yong and Pearce 2013) for both the PAF and ML methods, during which we found that the ML method does not converge properly and the loading coefficient of the PAF for the third factor for the geoid error is smaller than 0.3 (rule of thumb value for significant). As a result, the results for the 2-factor case computed by the PAF method are accepted. The unrotated and rotated loading coefficients, as well as the variance explained by each factor, are summarized in Table 3. Comparing them, we see that φ is more loaded into factor 1 whilst h is more correlated with factor 2 after the rotation. However, the other two variables are still looking like comprehensive ones, i.e. they are still significantly (>.3) contributing to more than one factor. Other oblique rotations give some better decompositions. However, the resulting factors are quite similar to the Varimax ones. At this stage, the results from the Varimax method are used in the following analysis.

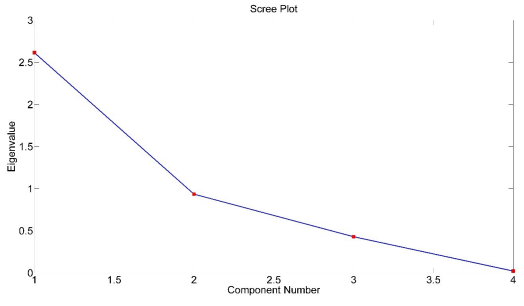

Figure 4: Scree plot of the FA analysis

The residuals of the 2-factor PAF model are shown in Fig. 5. The Q-Q plot of these residuals is shown in Fig. 6 also along with the counterparts from the linear models in Table 1 for the purpose of comparisons. From the Q-Q

Table 3: The unrotated and rotated loading coefficients of the PAF results from the 2-factor case.

PAF	Factors	φ	λ	h	δN	Variance Explained	Standard deviation of the residuals
Unrotated	1	0.82316	-0.79820	0.64002	0.94359	2.6147059	3.821 cm
	2	0.52213	0.37850	-0.65171	0.30672	0.9346838	
Rotated	**1**	**0.96818**	**-0.38121**	**0.08622**	**0.92593**	**1.9474762**	
	2	**0.11336**	**-0.79690**	**0.90935**	**0.35651**	**1.6019135**	

plot, it is clear that the FA method provides a better fit to the data. In addition, it also gives an estimation of the factor scores that may be useful during the search for some potential physical explanations of the NAVD 88 orthometric height errors. Fig. 7 shows the first factor of the model, which is basically describing the general trend from North-West to South-East. This agrees with some speculation of the accumulated systematic errors in the leveling network. However, to rigorously verify this, the original levelling routes that actually connect these benchmarks are needed to perform some simulation tests. For instance, one can add some predefined small errors in the actual leveling routes and propagate them into the orthometric heights at the NAVD 88 benchmarks. Then use these artificially distorted heights to compare with the gravimetric geoid to get the differences, based on which to repeat the FA analysis to see if the first factor corresponds to the accumulation effects to the leveling error. However, this original leveling information has not been prepared well for this purpose at this time though the author is eager to perform such kind of study. Thus, similar follow-up studies will be carried out once these resources are available. Fig. 8 shows the second factor that is highly correlated with the terrain, i.e. small in the flat states but large in the mountainous areas such as in the Rockies and the Appalachians. To make sure these terrain errors are not coming from the geoid model that suffers the downward continuation error inside of the masses from the topography to the geoid (Sjöberg 2007), the geopotential numbers from the NAVD 88 leveling and the geopotential number synthesized from the xGeoid16B reference model are compared directly on the surface of the Earth. These differences have similar patterns as the geoid differences. A quickly repeated FA analysis based on the geopotential differences shows the same thing as it has been described from Figs. 5-8. To save space, they are not plotted again here. Considering that the geopotential differences between two points are essentially the integration of the gravity values along the height increments, we can do a quick check on the gravity values used in NAVD 88 though currently we cannot repeat its leveling that is much more desired. As such, Fig. 9 shows the differences between the NAVD 88 gravity values and the XGeoid16B reference model predicted ones at these benchmarks (NAVD 88-XGeoid16refB). Figure 9 shows that the extreme values can reach almost up to 400 mGal. The size of the dots represents the magnitude of the difference. The color of the dots represents the sign. It is interesting to see that the NAVD 88 gravity tends to be systematically smaller than the XGeoid values in the mountainous areas, whereas it is systematically bigger in the flat areas.

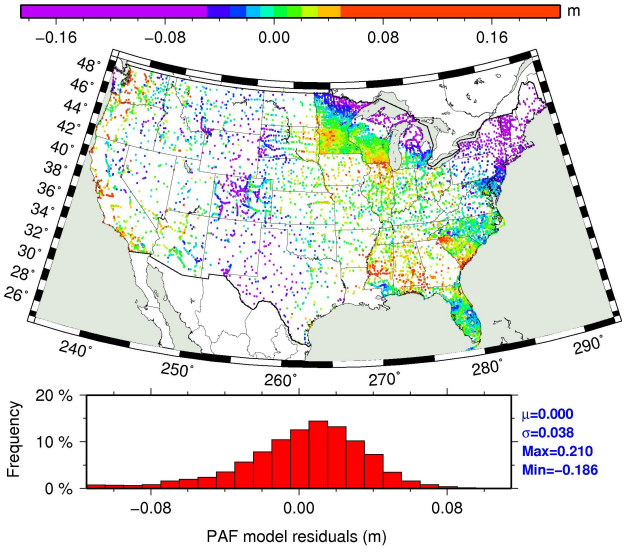

Figure 5: The model residuals of the PAF method (rms = 3.8 cm).

4.3 An extension of FA for predictions

The FA method is mainly used for analysis rather than prediction. However, for modeling purposes, it is necessary to predict the scores and the corresponding geoid differences at any location. This is fairly straightforward once we no-

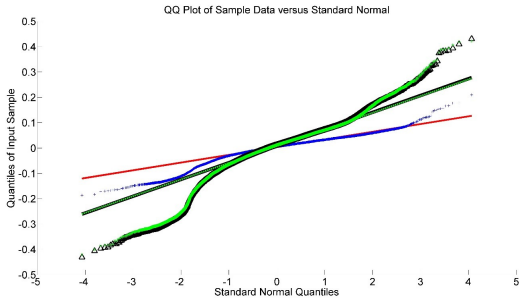

Figure 6: Q-Q plot of the FA residuals (blue signs and red line) vs. the counterparts both from the linear models $\tilde{y}= \beta_0^* + \beta_1^*\varphi + \beta_2^*\lambda + \epsilon$ (dark triangles and dark line) and $\tilde{y}= \beta_0^* + \beta_1^*\varphi + \beta_2^*\lambda + \beta_3^* h + \epsilon$ (green stars and green line) in Table 1.

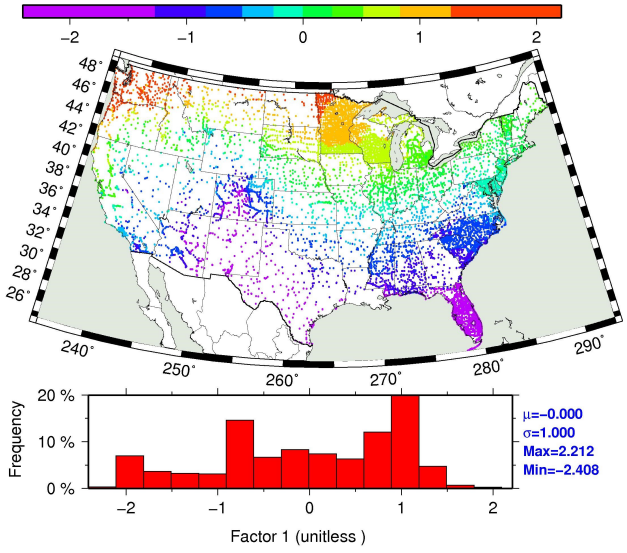

Figure 7: The first factor scores of the FA results.

tice the correlation between the observables, implied by the common factors $\vec{F}_i$, and the loading matrix W.

At any predicted point, the common factors are estimated based on the location of the data point and the already solved loading coefficients. Then, $\vec{X}_i$ is transposed into a column vector and separated into two parts:

$$\vec{y}_i^1 := \left[\varphi, \lambda, h^{'}\right]^T \tag{7}$$

and

$$\vec{y}_i^2 := [\delta N] \tag{8}$$

Without losing generality, Eq. (3) can be rewritten as:

$$\begin{bmatrix} \vec{y}_i^1 \\ \vec{y}_i^2 \end{bmatrix} = \vec{X}_i^T = W^T\vec{F}_i^T + \vec{\epsilon}_i^T = L\vec{F}_i^T + \vec{\epsilon}_i^T \tag{9}$$

with $L = W^T$.

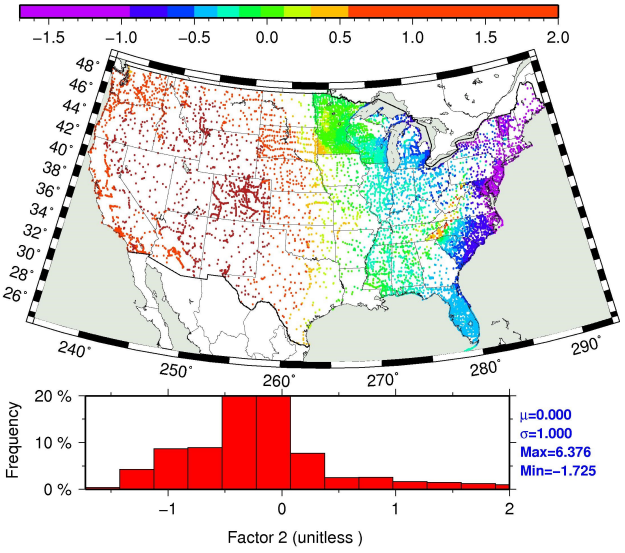

Figure 8: The second factor scores of the FA results.

The common factors $\vec{F}_i^T$ are estimated by solving the following system implied by $\vec{y}_i^1$ in Eq. (10)

$$\vec{y}_i^1 = L_1^{3\times 2}\vec{F}_i^T + \epsilon \tag{10}$$

Then the prediction is performed based on the solved common factors $\vec{F}_i^T$ from Eq. (10), as shown by Eq. (11)

$$\vec{y}_i^2 = L_2^{1\times 2}\vec{F}_i^T \tag{11}$$

where $L_2^{1\times 2}$ is the last row of W^T, and $L_1^{3\times 2}$ in Eq. (10) contains the first 3 rows of W^T.

One can understand the overall concept as using the first few rows in Eq. (10) to estimate $\vec{F}_i^T$ at this point and then using the last few rows to predict the desired values at the same place. It is clear that the loading matrix computed by FA plays an important role in the analysis, which opens a door other than LSC for making predictions of one kind of variables based on different kinds of observations at the same point but without using an analytical covariance model.

Considering that the actual prediction of the factors in Eq. (10) is only based on partial information contained in the full loading matrix, a rigorous mathematical expression and some practical numerical results should be given for the neglected terms. If all the information is used, under an equal weights assumption, the factor scores are given by:

$$\begin{aligned}\vec{F}_i^T &= \left[\begin{bmatrix} L_1 \\ L_2 \end{bmatrix}^T \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}\right]^{-1} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}^T \begin{bmatrix} \vec{y}_i^1 \\ \vec{y}_i^2 \end{bmatrix} \\ &= \left[L_1^T L_1\right]^{-1} L_1^T\vec{y}_i^1 - \frac{1}{1+g}\left[L_1^T L_1\right]^{-1}\left[L_2^T L_2\right]\left[L_1^T L_1\right]^{-1} L_1^T\vec{y}_i^1\end{aligned}$$

Table 4: Testing of the FA interpolation errors

Data used	Error Type	Prediction error (cm)	
		2 Factors	3 Factors
75%	Actual	6.92	6.18
	Formal	3.06	2.78

$$+\left[L_1^TL_1\right]^{-1}L_2^T\vec{y}_i^2-\frac{1}{1+g}\left[L_1^TL_1\right]^{-1}\left[L_2^TL_2\right]\left[L_1^TL_1\right]^{-1}L_2^T\vec{y}_i^2 \quad (12)$$

where $g = tr\left\{\left[L_2^TL_2\right]\left[L_1^TL_1\right]^{-1}\right\}$

The first term of the right hand side in Eq. (12) is the solution of the factor only based on the observation equation from Eq. (10). Thus, the extra terms in Eq. (12) are the prediction error of the factors, which is given by

$$\begin{aligned}\Theta_i = &-\frac{1}{1+g}L_2\left[L_1^TL_1\right]^{-1}\left[L_2^TL_2\right]\left[L_1^TL_1\right]^{-1}L_1^T\vec{y}_i^1\\&+L_2\left[L_1^TL_1\right]^{-1}L_2^T\vec{y}_i^2\\&-\frac{1}{1+g}L_2\left[L_1^TL_1\right]^{-1}\left[L_2^TL_2\right]\left[L_1^TL_1\right]^{-1}L_2^T\vec{y}_i^2\end{aligned} \quad (13)$$

Because the FA separates the specific errors ψ in the data from the loading matrix W, the formal prediction error in Eq. (14) does not contain the observation errors in the data. As such, it is usually smaller than the actual misfits that contains random observation errors. For instance, 75% of the data in Fig. 1 are sampled (Green 1977) as the control to estimate the loading matrix, while the other 25% are used to check the differences between the FA predicted values.

The precision of the FA predicted value is 6.92 cm, while the standard deviation of the formal errors computed from Eq. (14) is only 3.06 cm. The differences are due to neglecting of the specific error ψ in Eq. (11) and the least square error during estimating the common factors in Eq. (10). If 3 Factors are used, the formal error is reduced from 3.06 cm into 2.78 cm, and the prediction error is also reduced from 6.92 cm into 6.18 cm. This reduction is mainly due to the power transition from the specific errors into systematic effects by using one extra factor. All the prediction precisions are summarized in Table 4.

5 Evaluations

The NAVD 88 orthometric height errors modeled by all of the above methods were removed from the actual orthometric height values at all of the 21,112 stations that were used in the last NGS hybrid geoid model, i.e. Geoid12B, which accordingly gives several versions of calibrated NAVD 88 heights. To investigate which method yields the best results, some of the recently published gravity field models from ICGEM are used as external reference. Note, to try to see the differences at various resolutions, not only the higher degree and order models but also some of the satellite only models are used here. Please see more detailed descriptions of these models (such as data sources, data combination schemes, solution technology used, and model resolutions) at http://icgem.gfz-potsdam.de/home. First, on each GPS/Leveling benchmark, the original orthometric heights are calibrated by removing the systematic errors modeled by all the above mentioned methods such as the linear model, the cubic spline model, the Legendre polynomial model and the FA model. Then, several versions of geoid heights are obtained based on these newly calibrated orthometric heights. Finally, the geoid undulations computed from the ICGEM models are compared with these geoid heights based on these calibrated orthometric heights. The logarithm of the standard deviations of the differences are shown in Fig. 10, where the horizontal axis is ordered by the maximum degree of the corresponding models, from low to high from left to right. Before using any calibration of the NAVD 88 errors, the model precision changes from low degrees to high degrees are relatively small, as indicated by the heights of the dark blue bars from left to right in Fig. 10. After removing the systematic errors modeled by the linear model (cyan bars), the B-splines (yellow bars), the Legendre functions (red bars), and FA (brown bars), the degree effects from various models can be identified more clearly, especially for the FA results. For the low degree models starting from Tongji-grace01 (Chen et al. 2015) to DIR5 (Bruinsma et al. 2014), the differences between different error modeling techniques are not substantial, though they all tend to agree better to the calibrated heights except one GRACE only model. Starting from the GGM05c model, it is clear that the FA method consistently provides better agreements with the ICGEM models. This is even clear for the recent models such as GECO and EIGEN6c4, which may be due to the recently new applied techniques in these newer models.

6 Conclusions

The differences between the geoid model-implied undulations and the ground observed geoid heights at over 20,000 GPS/Leveling benchmarks in the CONUS area were

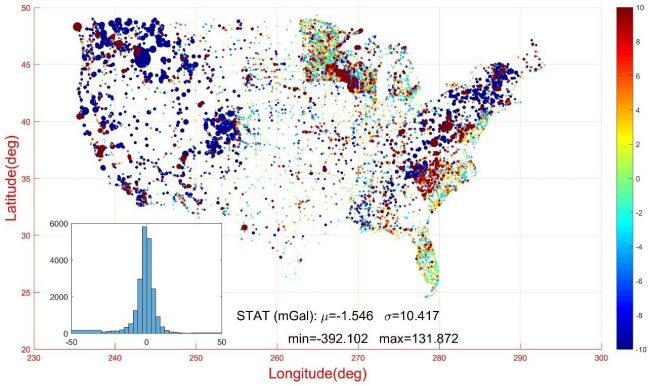

Figure 9: The gravity differences between NAVD 88 and XGeoid16refB at the NAVD 88 benchmarks (the size of the dot represents the magnitude with positive towards red and negative towards blue).

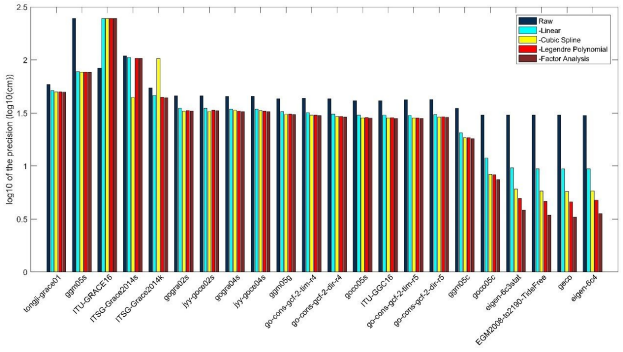

Figure 10: The logarithm plot of the precision of various ICGEM models at the NAVD 88 benchmarks and all the counterparts after removing the systematic errors modeled by a linear model, a B-Spline model, a Legendre-polynomial model, and FA model.

used as data samples for determining the NAVD 88 errors in this continental region. Both the standard regression analyses and the FA method were applied to analyze and comprehend these errors. The normality plots of the residuals show that the NAVD 88 error embedded in the orthometric heights is not characterized only by a simple "tilt" as widely used in many previous studies. By comparing the results and the corresponding analysis steps, it is clear that the FA algorithm produced more accurate results with the added advantage of being, in the author's opinion, an efficient and elegant process. The common steps of selecting or defining the base functions and the associated time consuming trial and error stage in the regression analysis could be avoided.

In addition to giving a relatively compact representation of the errors, the factor scores from FA provided some plausible explanations of the underlying causes creating the NAVD 88 error. Further studies are still required to absolutely verify them once enough information of the leveling is obtained. In addition to the analysis of the control data, a predictive methodology was developed by using the relationship between the loading matrix and the common factors. Based on the estimated loading coefficients, the factor scores at any given location can be computed, and then used for predictive purposes. This novel usage opens a new door to make predictions of one kind of variables based on the measurements of other kind of variables without using an analytical covariance model that is usually used in the LSC. The prediction error analysis tells that after removing the systematic errors in NAVD 88, one can expect 3-7 cm random errors, which is especially useful for evaluating the vast amount of historical applications based on NAVD 88, and provides error budgets for stakeholders to make mitigation decisions.

Finally, independent geoid models from ICGEM were used to validate the different methodologies. The FA method provided the best fits to all ICGEM models, especially to the newer ones where new technologies and data have been used during generating these models. The better agreements show the improvements of these newer models more clearly.

Abbreviations

CONUS The conterminous USA
FA Factor Analysis
GOCE Gravity and Ocean Circulation Explorer
GRACE Gravity Recovery and Climate Experiment
GRAV-D Gravity for the Redefinition of the American Vertical Datum
ICGEM International Centre for Global Earth Models
LSC Least Squares Collocation
LVA Latent Variable Analysis
NAVD 88 North American Vertical Datum of 1988
NGS National Geodetic Survey
PAF Principal Axis Factoring
PCA Principal Component Analysis
PDF Probability Density Function
Q-Q Quantile by Quantile

Acknowledgement: The author thanks for the useful comments and suggestions from reviewer 1 and reviewer 2 (Dr. Elena Rangelova at University of Calgary).

A The Computational Details of Using B-splines and Legendre Functions

In addition to the polynomial functions, some other local base functions based on B-splines and Legendre functions are also used to model the NAVD 88 errors. The corresponding mathematical expressions for B-splines and Legendre functions are described in Eq. (A1) and Eq. (A6), respectively.

$$\tilde{\zeta}_{x=\varphi/\Delta,y=\lambda/\Delta} := \sum_{k=0}^{3}\sum_{l=0}^{3} \alpha_{(i+k)(j+l)} B_k(s) B_l(t) + \epsilon, \qquad \text{(A1)}$$

where $\alpha_{(i+k)(j+l)}$ is the coefficient on the control lattice, $i = \lfloor x \rfloor - 1$, $j = \lfloor y \rfloor - 1$, $s = x - \lfloor x \rfloor$, $t = y - \lfloor y \rfloor$, $x = \varphi/\Delta$, $y = \lambda/\Delta$, Δ is the resolution of the control lattice. B_k and B_l are uniform B-spline basis functions defined as:

$$B_0(t) = (1-t)^3/6, \qquad \text{(A2)}$$

$$B_1(t) = (3t^3 - 6t^2 + 4)/6, \qquad \text{(A3)}$$

$$B_2(t) = (-3t^3 + 3t^2 + 3t + 1)/6, \qquad \text{(A4)}$$

$$B_3(t) = t^3/6. \qquad \text{(A5)}$$

The regression model based on the Legendre functions reads as:

$$\tilde{\zeta}(\vec{r}) := \sum_{i=1}^{I} \alpha_i \left\{ \sum_{n=n1}^{n2} P_n \left(\frac{\vec{r}}{|\vec{r}|} \cdot \vec{R}_i \right) \right\} + \epsilon, \qquad \text{(A6)}$$

where α_i is the coefficient on each of the control points $(1, \cdots, I)$ located at $\vec{R}_i$, $\vec{r}$ is the vector of the observation point, and P_n is the nth-degree Legendre polynomial, $n1$ and $n2$ are the lower and upper spectrum limits of the model; please see Eicker (2008), among others, for the details on using this kind of localized functions.

Figs. A1-A2 show the standard deviations of the fitting residuals versus the number of parameters that are needed for the B-splines and the Legendre polynomials, respectively. One can use them to roughly determine the number of parameters that are needed to model the data, while, at the same time, avoiding extreme interpolation errors between points caused by over parameterization. For instance, Fig. A2 shows that only 6 parameters are sufficient to capture the NAVD 88 error in the band of degree 0 to degree 4 when Eq. (A6) is applied. The standard deviation of the fitting residual is about 5 cm. Fig. A1 shows that many more parameters will be required to achieve about the same results if the B-spline base functions are used.

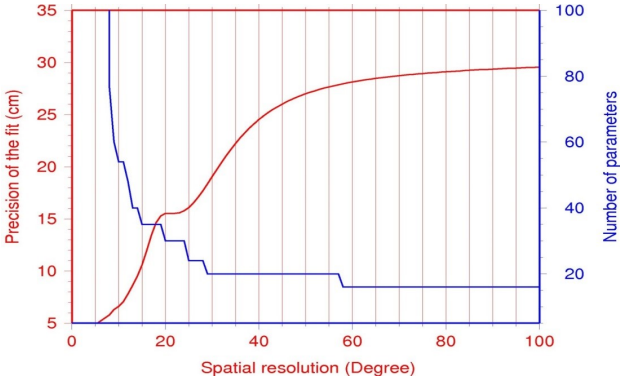

Figure A.1: The model precision and the number of the corresponding coefficients versus the spatial resolution of the B-spline model. The red curve is the model precision referring to the left axis. The blue curve is the number of the parameters referring to the right.

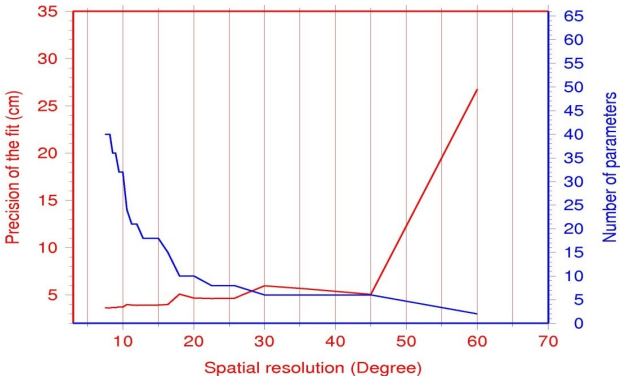

Figure A.2: The model precision and the number of the corresponding coefficients versus the spatial resolution of the Legendre polynomial model. The red curve is the model precision referring to the left axis. The blue curve is the number of the parameters referring to the right.

B The Derivation of Solving the FA Problem with the PAF Method and the ML Method

B.1 Principal Axis Factoring (PAF)

Starting from Eq. (10), we perform a linear regression of each variable j on all other variables and set ψ_j to the mean square error for that regression so that we obtain an estimation of ψ. Then we define the reduced covariance matrix U as:

$$U = V - \psi. \qquad \text{(B1)}$$

Eq. (B1) shows how much of the variance in each variable is associated with the variances of the latent factors (Shalizi 2009). Because U is a symmetric matrix, there is a

spectral decomposition of it, as shown in Eq. (B2).

$$U = \left(C_q D_q^{1/2}\right)\left(C_q D_q^{1/2}\right)^T, \tag{B2}$$

where C_q is the matrix whose columns are the eigenvectors of U, and $D_q^{1/2}$ is a diagonal matrix of the square roots of the eigenvalues of U. Then we obtain an estimate of W as:

$$W = \left(C_q D_q^{1/2}\right)^T. \tag{B3}$$

Eq. (B3) is used to re-set the specific errors according to Eq. (B4).

$$\psi_j = V_{jj} - \sum_{r=1}^{q} W_{rj}^2. \tag{B4}$$

Then the iteration starts until some threshold is met. It is known that the convergence is very quick and does not depend very much on the accuracy of the first estimate of the specific errors (Shalizi 2009), as long as they are not too extreme.

B.2 ML

If $\vec{X}_i$ is assumed to have multi-normal distributions, the likelihood function of $\vec{X}_i$ is given by:

$$lik\left(\vec{X}_i; \Sigma = \psi + W^T W\right) = (2\pi)^{-p/2}|\Sigma|^{-1/2} exp\{-\frac{1}{2}\vec{X}_i^T \Sigma^{-1}\vec{X}_i\}. \tag{B5}$$

The overall sample likelihood function reads as:

$$lik\left(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n; \Sigma = \psi + W^T W\right) = (2\pi)^{-np/2}|\Sigma|^{-n/2} exp\{\sum_{i=1}^{n}(-\frac{1}{2}\vec{X}_i^T \Sigma^{-1}\vec{X}_i)\}. \tag{B6}$$

The log of the likelihood function is:

$$l = \log\left\{lik\left(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n; \Sigma = \psi + W^T W\right)\right\} = -\frac{np}{2}\log(2\pi) - \frac{n}{2} log\,|\Sigma| - \sum_{i=1}^{n}(-\frac{1}{2}\vec{X}_i^T \Sigma \vec{X}_i). \tag{B7}$$

The last term in the log likelihood function is:

$$\sum_{i=1}^{n}(\vec{X}_i^T \Sigma^{-1}\vec{X}_i) = \sum_{i=1}^{n} tr(\vec{X}_i^T \Sigma^{-1}\vec{X}_i) = n\sum_{i=1}^{n} tr(\frac{\vec{X}_i \vec{X}_i^T}{n}\Sigma^{-1}) = n \cdot\ tr\left(\sum_{i=1}^{n}\frac{\vec{X}_i \vec{X}_i^T}{n}\Sigma^{-1}\right) = n \cdot tr(V\Sigma^{-1}). \tag{B8}$$

Combining Eqs. (a7-a8) gives the final log likelihood function as shown in Eq. (B9).

$$l = -\frac{np}{2}\log 2\pi - \frac{n}{2}\log\left|\psi + W^T W\right| - \frac{n}{2} tr\left(V(\psi + W^T W)^{-1}\right) \tag{B9}$$

In theory, one can set the partial derivatives of l with respect to the parameters in Σ to zero and solve the system. But in practice, the following alternative target function T_{ML} in Eq. (B10) is used, after noticing that minimizing T_{ML} is essentially equivalent to maximizing l (Joreskog 1969, Bartholomew 1987, Shalizi 2009). Note that the biased sample variance is replaced with the unbiased estimates here.

$$T_{ML} := \log|\psi + W^T W| + tr\left((\psi + W^T W)^{-1} V\right) - \log|V| - p. \tag{B10}$$

Similar numerical iterations are normally used to achieve the (numerical) minimum value of T_{ML}. The differences between the ML and the PAF are normally not very substantial. However, the Cramér–Rao theorem predicts that when the variables are multi-normal and the common factor model holds in the population, the ML method generates the solution that most accurately reflects the underlying population pattern (de Winter and Dodou 2012).

References

Bartholomew D.J., 1987, Latent Variable Models and Factor Analysis, New York: Oxford University Press.

Browne M.W., 2001, An overview of analytic rotation in exploratory factor analysis, Multi. Behav. Res., 36, 111-150, doi: https://doi.org/10.1207/S15327906MBR3601_05.

Bruinsma S., Förste C., Abrikosov O., Lemoine J.M., Marty J.C., Mulet S., Rio M.H., and Bonvalot S., 2014, ESA's satellite-only gravity field model via the direct approach based on all GOCE data, Geophys. Res. Lett., 41, 21, 7508–7514, doi: 10.1002/2014GL062045.

Cattell R.B., 1965, Factor Analysis: An Introduction to Essentials I. The Purpose and Underlying Models, Biometrics, 21, 1, 190-215, doi: 10.2307/2528364.

Cawley G.C., and Talbot N.L.C., 2010, On over-fitting in model selection and subsequent selection bias in performance evaluation. J Mach Learn Res., 11, 2079–2107, http://dblp.uni-trier.de/db/journals/jmlr/jmlr11.html#CawleyT10.

Chen Q., Shen Y., Zhang X., Hsu H., and Chen W., 2015, Global Earth's gravity field solution with GRACE orbit and range measurements using modified short arc approach, Acta Geod Geophys; 50, 2,173-185, doi: https://doi.org/10.1007/s40328-014-0077-1.

Eicker A., 2008, Gravity field refinements by radial basis functions from in-situ satellite data, Ph.D. dissertation, Univ. of Bonn, Bonn, Germany. (Available at http://hss.ulb.uni-bonn.de/diss_online/landw_fak/2008/eicker_annette).

Featherstone W.E., and Filmer M.S., 2012, The north-south tilt in the Australian Height Datum is explained by the ocean's

mean dynamic topography, J. Geophys. Res., 117, C08035, doi: 10.1029/2012JC007974.

Förste C., Bruinsma S., Abrykosov O., Flechtner F., Marty J.C., Lemoine J.M., Dahle C., Neumayer K.H., Barthelmes F., König R., and Biancale R., 2014, EIGEN-6C4 - The latest combined global gravity field model including GOCE data up to degree and order 1949 of GFZ Potsdam and GRGS Toulouse, (Geophysical Research Abstracts, Vol. 16, EGU2014-3707, 2014), General Assembly European Geosciences Union.

Gruber T., 2014, Scientific roadmap towards height system unification with GOCE, the GOCE5 user workshop, www.iapg.bgu.tum.de/mediadb/7184044/.../20141127_Gruber_GOCE_HSU.pdf.

Green B.F.J., 1977, FORTRAN subroutines for random sampling without replacement. Behavior Research Methods & Instrumentation, 9, 6, 559-559, doi: https://doi.org/10.3758/BF03214009.

Hocking R.R., 1976, The Analysis and Selection of Variables in Linear Regression, Biometrics, 32, 1, 1-49, doi: 10.2307/2529336.

Joreskog K.G., 1969, A General Approach to Confirmatory Maximum Likelihood Factor Analysis, Psychometrika, 34, 2, 183-202, doi: https://doi.org/10.1007/BF02289343.

Kaiser H.F., 1958, The Varimax criterion for analytic rotation in factor analysis, Psychometrika, 23, 3, 187-200, doi: https://doi.org/10.1007/BF02289233.

Klees R., and Prutkin I., 2010, The combination of GNSS-levelling data and gravimetric (quasi-) geoid heights in the presence of noise, J Geod., 84, 12, 731-749, doi: https://doi.org/10.1007/s00190-010-0406-2.

Koch K.R., 1988, Parameter Estimation and Hypothesis Testing in Linear Models. Springer, New York.

Li X., Crowley J.W., Holmes S.A., and Wang Y.M., 2016, The contribution of the GRAV-D airborne gravity to geoid determination in the Great Lakes region, Geophys. Res. Lett., 43, 9, 4358–4365, doi: 10.1002/2016GL068374.

Li X., 2017, Using radial basis functions in airborne gravimetry for local geoid improvement, J Geod., https://doi.org/10.1007/s00190-017-1074-2,

Mulaik S.A., 1972, The Foundations of Factor Analysis, New York: McGraw-Hill Book Co.

Pavlis N.K., Holmes S.A., Kenyon S.C., and Factor J.K., 2012, The development and evaluation of the Earth Gravitational Model 2008 (EGM2008), J. Geophys. Res. 117, B04406, doi:10.1029/2011JB008916.

Prutkin I., and Klees R., 2008, On the non-uniqueness of local quasi-geoids computed from terrestrial gravity anomalies, J Geod., 82, 3, 147-156, doi: https://doi.org/10.1007/s00190-007-0161-1.

Rencher A.C., 2002, Methods of Multivariate Analysis. John Wiley and Sons, New York.

Roman D.R., Wang Y.M., Henning W., and Hamilton J., 2004, Assessment of the new national geoid height model, GEOID03. In: Proceedings of the American Congress on Surveying and Mapping 2004 meeting.

Roman D.R., Wang Y.M., Saleh J., and Li, X., 2010a, Geodesy, geoids, and vertical datums: A perspective from the U.S. National Geodetic Survey. In: Facing the Challenges Building the Capacity, FIG Congress 2010, Sydney, 1-16.

Roman D.R., Wang Y.M., Saleh J., and Li X., 2010b, A Gravimetric Geoid Model for the United States: The Development and Evaluation of USGG2009, National Geodetic Survey, National Ocean Service, National Oceanic and Atmospheric Administration, US Department of Commerce, Silver Spring, MD. http://www.geodesy.noaa.gov/GEOID/USGG2009/USGG2009_tech_details.pdf

Roman D.R., Wang Y.M., Saleh .J, and Li X., 2010c, Final National Models for the United States: Development of GEOID09, National Geodetic Survey, National Ocean Service, National Oceanic and Atmospheric Administration, US Department of Commerce, Silver Spring, MD. www.geodesy.noaa.gov/GEOID/GEOID09/GEOID09_tech_details.pdf.

Roman D.R., and Li X., 2014, GRAV-D: Using Aerogravity to Produce a Refined Vertical Datum, Article of the Month, September 2014, FIG congress 2014, https://www.fig.net/pub/monthly_articles/september_2014/roman_li.html.

Rummel R.J., 1988, Applied Factor Analysis, Northwestern University Press.

Rummel R., Yi W., and Stummer C., 2011, GOCE gravitational gradiometry, J Geod., 85, 11, 777–790, doi: 10.1007/s00190-011-0500-0.

Rummel R., 2012, Height unification using GOCE, J. Geod. Sci., 2, 4, 355–362, doi: https://doi.org/10.2478/v10156-011-0047-2.

Shalizi C., 2009, Lecture notes from Carnegie Mellon University, 2009, http://www.stat.cmu.edu/~cshalizi/350/lectures/12/lecture-12.pdf

Smith D., 2007, The GRAV-D project: Gravity for the Redefinition of the American Vertical Datum. Available online at: http://www.ngs.noaa.gov/GRAV-D/pubs/GRAV-D_v2007_12_19.pdf.

Smith D.A., Holmes S.A., Li X., Guillaume Y., Wang Y.M., Bürki B., Roman D.R., and Damiani T.M., 2013, Confirming regional 1 cm differential geoid accuracy from airborne gravimetry: the Geoid Slope Validation Survey of 2011, J Geod., 87, 10–12: 885–907. doi: http://dx.doi.org/10.1007/s00190-013-0653-0.

Snedecor G.W., and Cochran W.G., 1989, Statistical Methods, Eighth Edition, Iowa State University Press. ISBN 978-0-8138-1561-9.

Sjöberg L.E., 2007, The topographic bias by analytical continuation in physical geodesy, J Geod., 81, 5, 345-350, doi: https://doi.org/10.1007/s00190-006-0112-2.

Suhr D.D., 2005, Principal Component Analysis vs. Exploratory Factor Analysis. Proceedings of the 30th Annual SAS Users Group International Conferences, (UGIC'05), SAS Institute Inc., 1-157.

Tapley B.D., Bettadpur S., Watkins M.M., and Reigber C., 2004, The gravity recovery and climate experiment: Mission overview and early results. Geophys Res Lett, 31 doi: 10.1029/2004GL019920.

de Winter J.C..F, and Dodou D., 2012, Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. Journal of Applied Statistics Vol. 39, 4, 695–710.

Yong A.G., and Pearce S.A., 2013, Beginner's guide to factor analysis: focusing on exploratory factor analysis, Tutor Quant Methods Psychol., 9, 2, 79–94, doi: DOI: 10.20982/tqmp.09.2.p079.

Zilkoski D., Richards J., and Young G., 1992, Results of the General Adjustment of the North American Vertical Datum of 1988, Surveying and Land Information Systems, 52, 3, 133–149.