（3）筆界復元はなぜ座標変換なのか

データ解析の基本は大量のデータを使ってあらゆるパターンで証明することが基本です，以前，少なくともウインドウズ95が普及する1995年以前は少ないデータで説明されたものが多く，大量のデータでは説明出来ていないものもありましたがパソコンの普及により大量のデータ処理が可能となりました。

筆界復元はなぜ座標変換なのか・その 1，その2では294個のデータを誤差の大きさで4パターンに点の位置条件で30組程度のデータ解析結果で説明してきました，紙面の都合で全ての結果をお見せしたわけではありませんがおおよその説明は出来たと思っています。この後も同じデータで最低30組程度のデータで計算，解析してみます。

変数減少法について

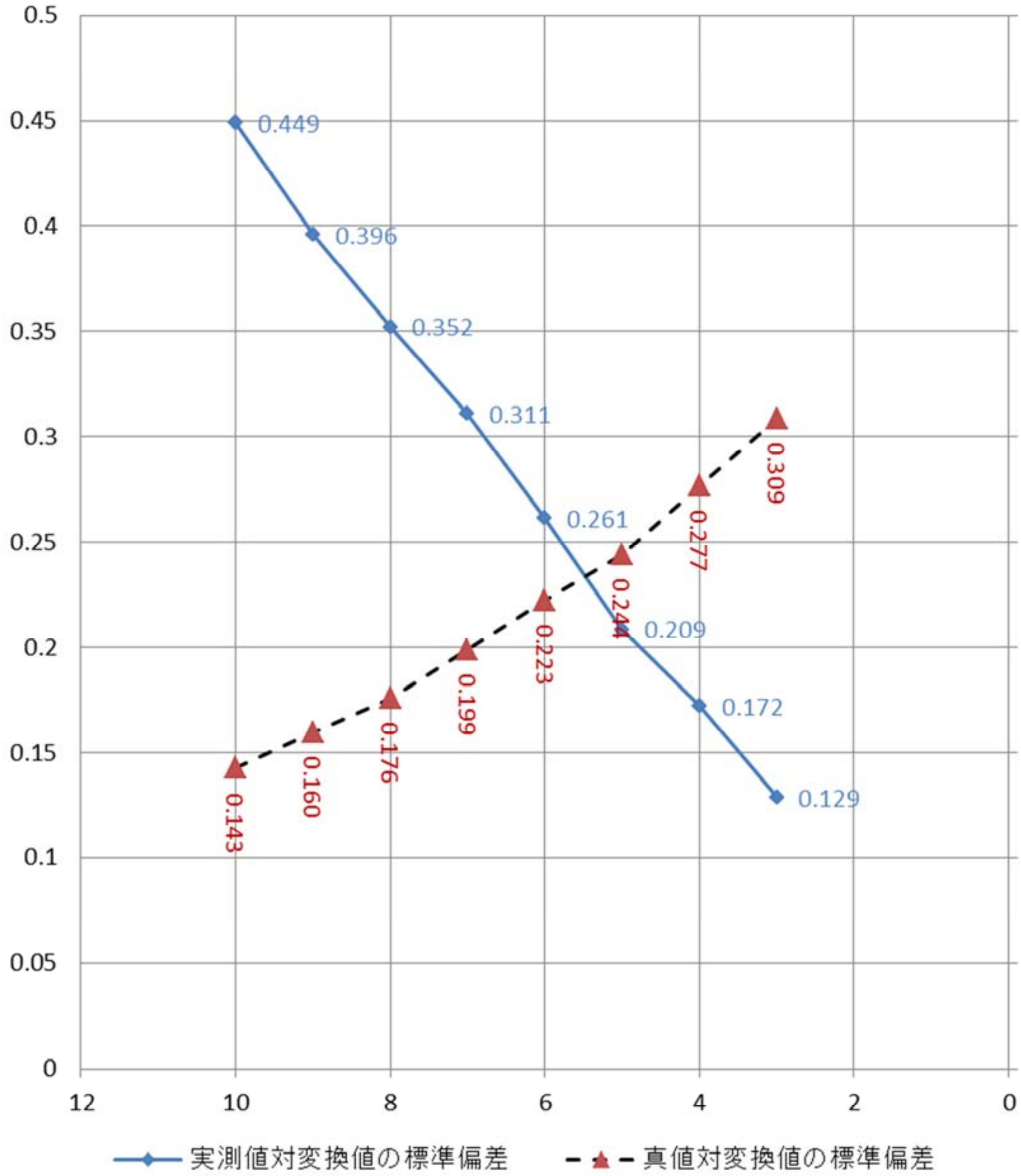
座標変換で境界復元をするには出来るだけ多くの境界標を準拠点（計算の基点にする点）に使って座標変換すれば良いことは知られています，しかもその座標変換方法は最小二乗法座標変換であることが条件です。最小二乗法に関しては「別のページ」で説明するとしてここでは最小二乗法座標変換のなかで一般的なヘルマート変換を使って確認しています。

変数減少法とは図面値対実測値で ①最小二乗法座標変換し標準偏差を記録，②実測値と変換値の差が大きいデータを削除，③手順①と②を繰り返す，という方法です。通常は手順①の結果を得て，実測値と変換値の差の大きい順から小さい順（逆順）に並べれば判ることです，つまり大きい順に並べたらg1，g2，g3，g4，・・・・g10となっているデータを大きい方から3個除くとg1，g2，g3となるのですが，最小二乗法座標変換を間に挟むことによりこの順番どおりにならない事があります，これは組合せによって関係位置が変わるからです，つまり順番がg1，g3，g4になることがあります，その結果g2は2番目に悪いデータでは無いということになります。

実測値と変換値の差が大きいデータ順に削除していきますので繰り返す毎に標準偏差は小さくなります。このときに真値と変換値の差（較差）の標準偏差も計算しておきます，この標準偏差の一番小さい状態が最も精度（標準偏差）の出る点の組合せになります。ランダムにデータを選んで変数減少法により得られたデータを縦軸に実測値と変換値の差（較差）の標準偏差と真値と変換値の差（較差）の標準偏差を横軸に準拠点数をグラフにしました。感覚的に実測値と変換値の差（較差）の標準偏差が小さくなれば復元精度（標準偏差）も小さくなると思いがちですが実測値と変換値の差（較差）の標準偏差と復元精度（標準偏差）は比例しないということです。

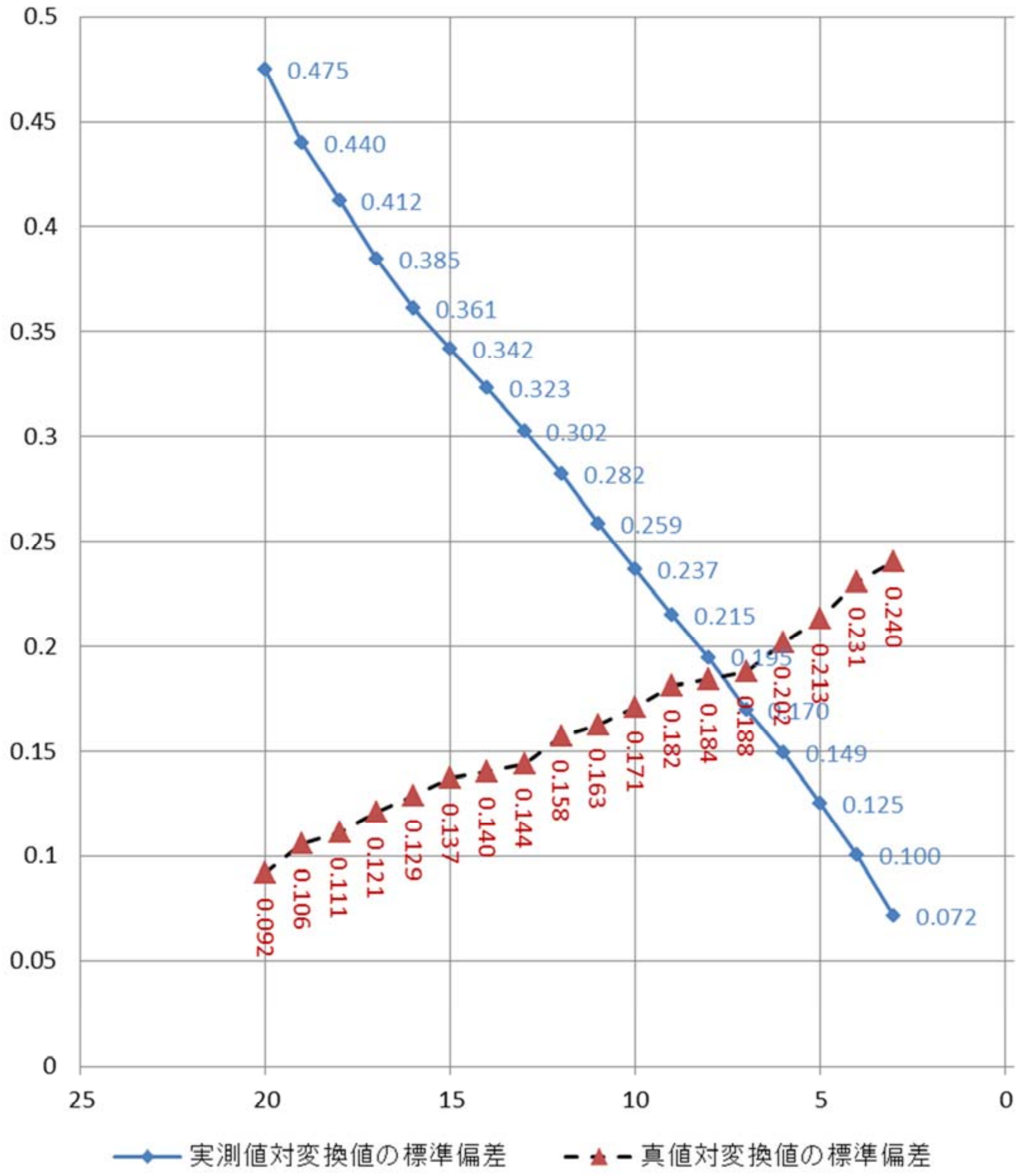
与えてある誤差は標準偏差ベースで0．487，準拠点数10個～3個，30回の平均値

ここでは正規分布から外れる誤差は入っていませんので実測値対変換値の標準偏差が与えた標準偏差ベースの誤差0．487に近いこと(点数の補正をして)が必要です。グラフでは10点で0．449ですから補正値1．07を乗じて0．480で与えた標準偏差ベースの誤差0．487に近いことが分かります。与えた標準偏差ベースの誤差0．487ですから10点目の0．143で見れば3．4倍の精度が得られているということになります。



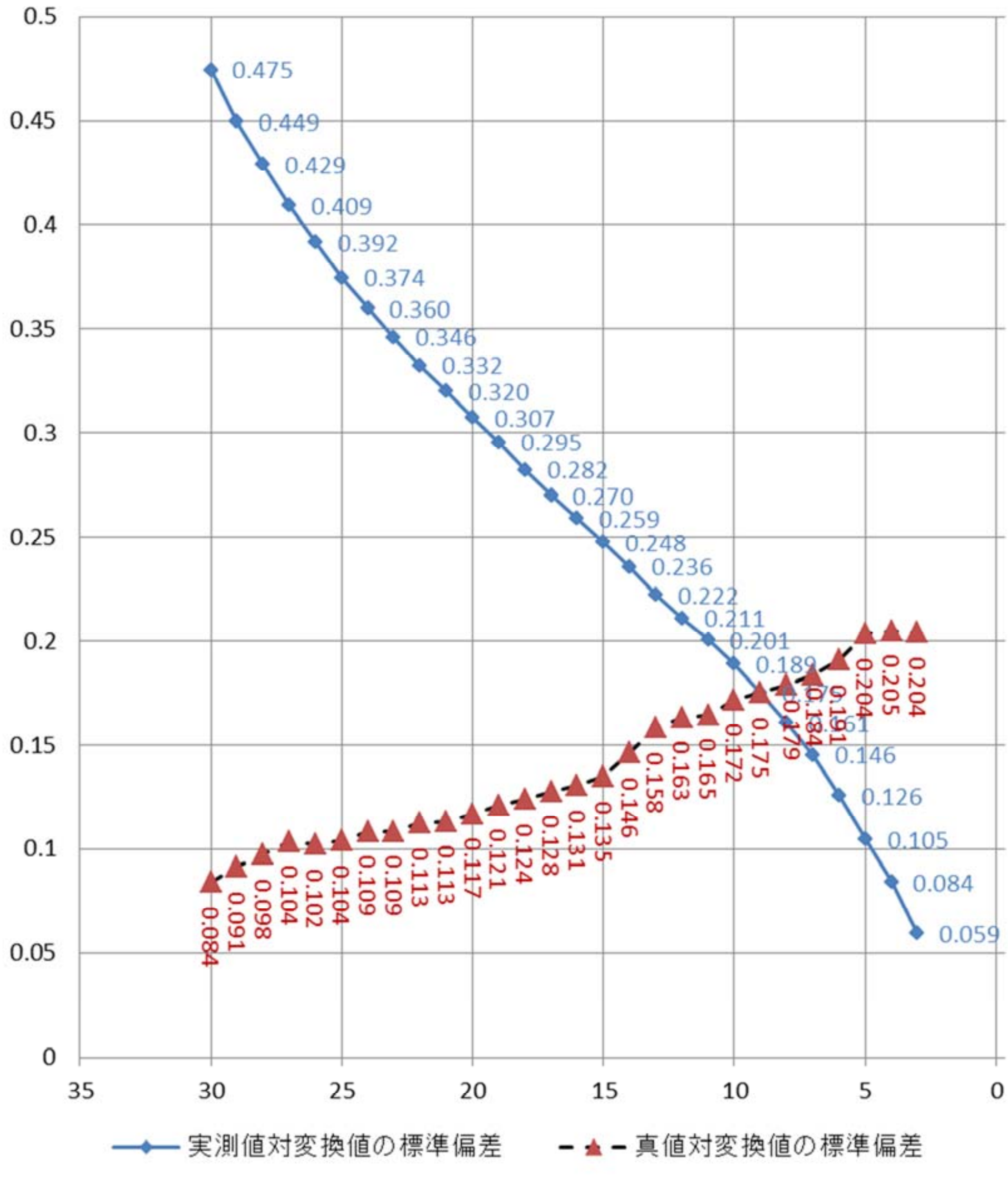
与えた標準偏差ベースの誤差0．487，準拠点数20個～3個，30回の平均値

グラフでは20点で0．475ですから補正値1．04を乗じて0．494で与えた標準偏差ベースの誤差0．487に近いことが分かります。与えた標準偏差ベースの誤差0．487ですから20点目の0．092で見れば5．3倍の精度が得られているということになります。



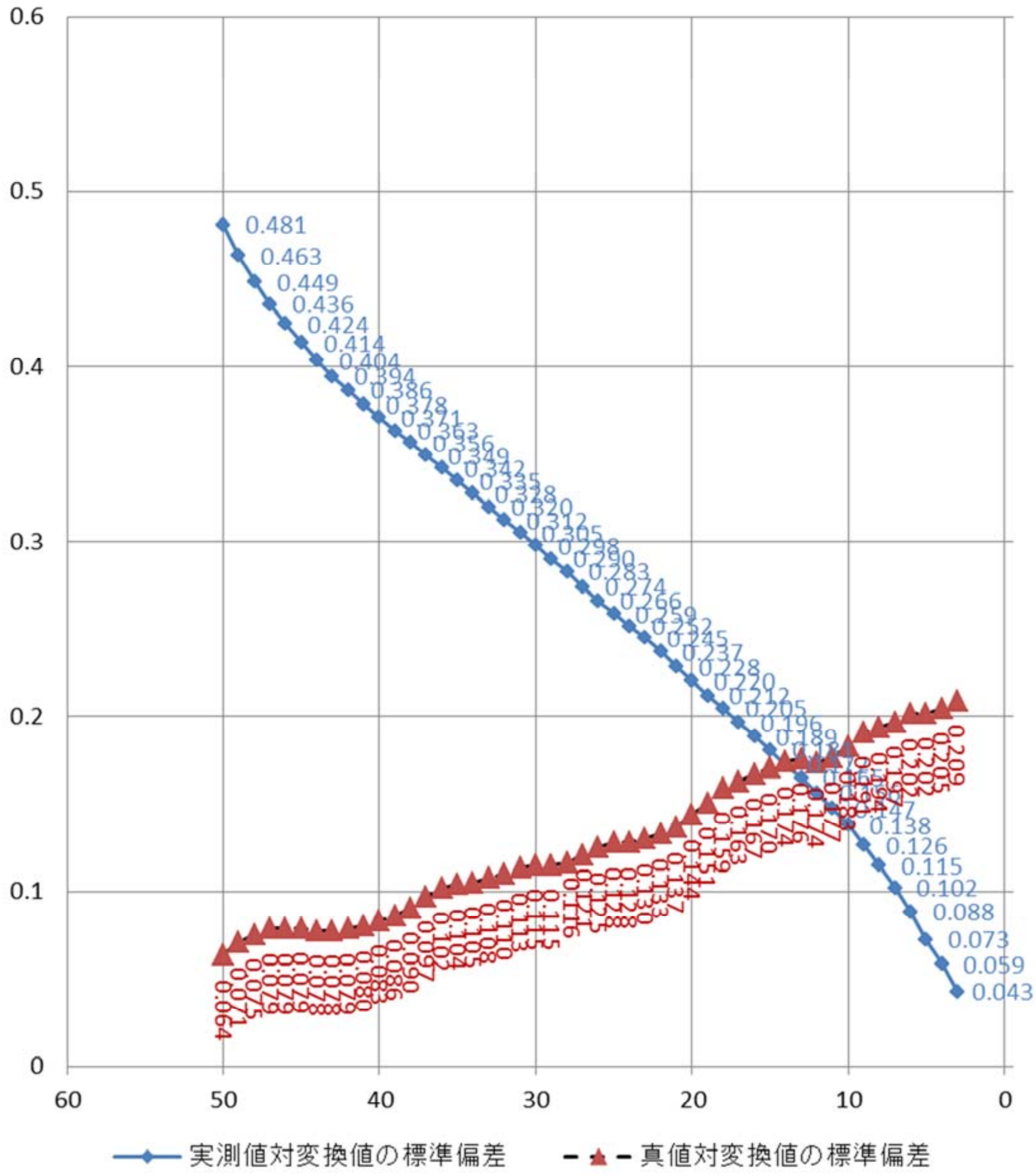
与えた標準偏差ベースの誤差で0．487，準拠点数30個～3個，30回の平均値

グラフでは30点で0．475ですから補正値1．02を乗じて0．485で与えた標準偏差ベースの誤差0．487に近いことが分かります。与えた標準偏差ベースの誤差0．487ですから30点目の0．084で見れば5．8倍の精度が得られているということになります。



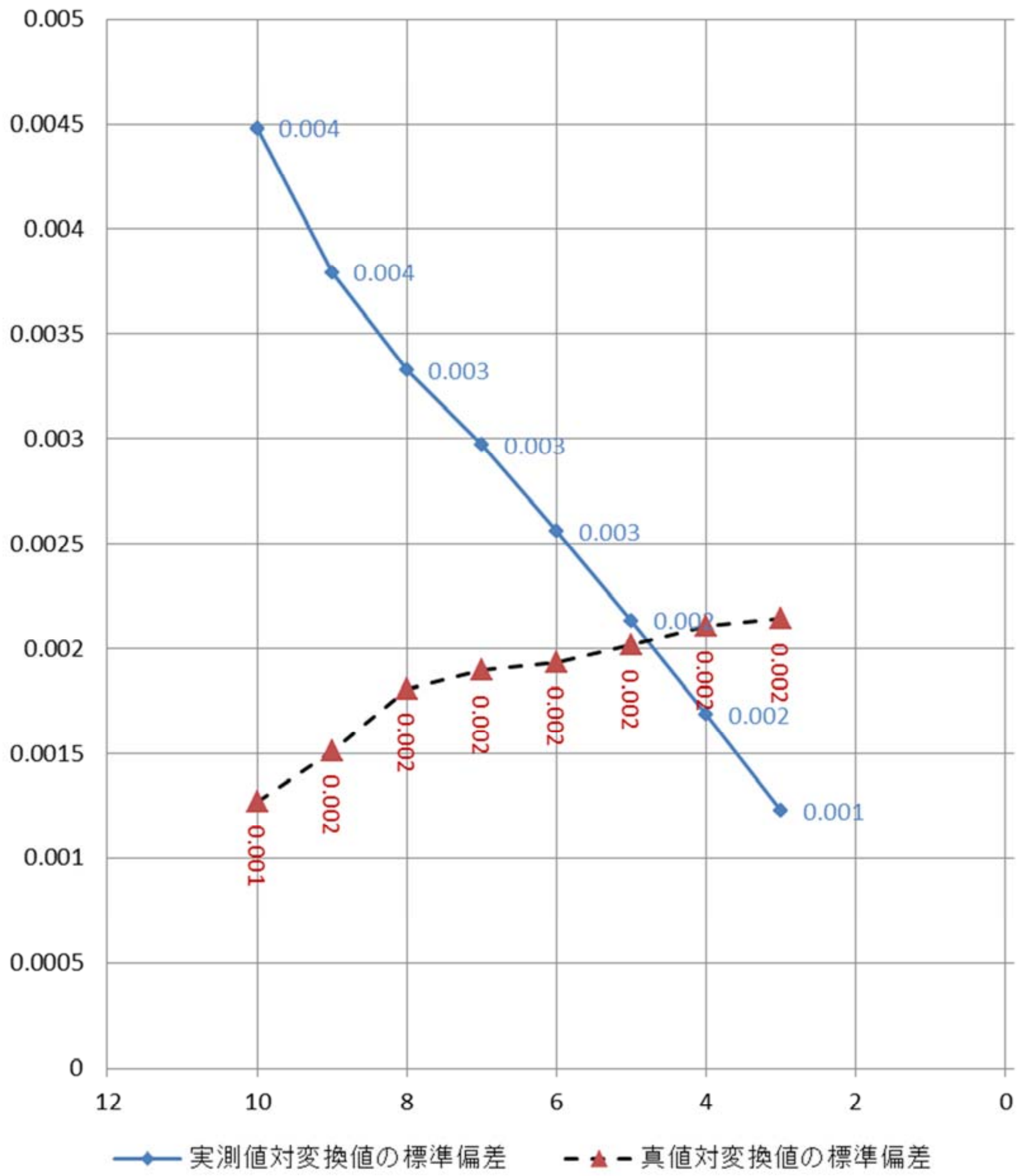
与えてある誤差は標準偏差ベースで0．487，準拠点数50個～3個，30回の平均値

グラフでは50点で0．481ですから補正値1．00を乗じて0．481で与えた標準偏差ベースの誤差0．487に近いことが分かります。与えた標準偏差ベースの誤差0．487ですから50点目の0．064で見れば7．6倍の精度が得られているということになります。

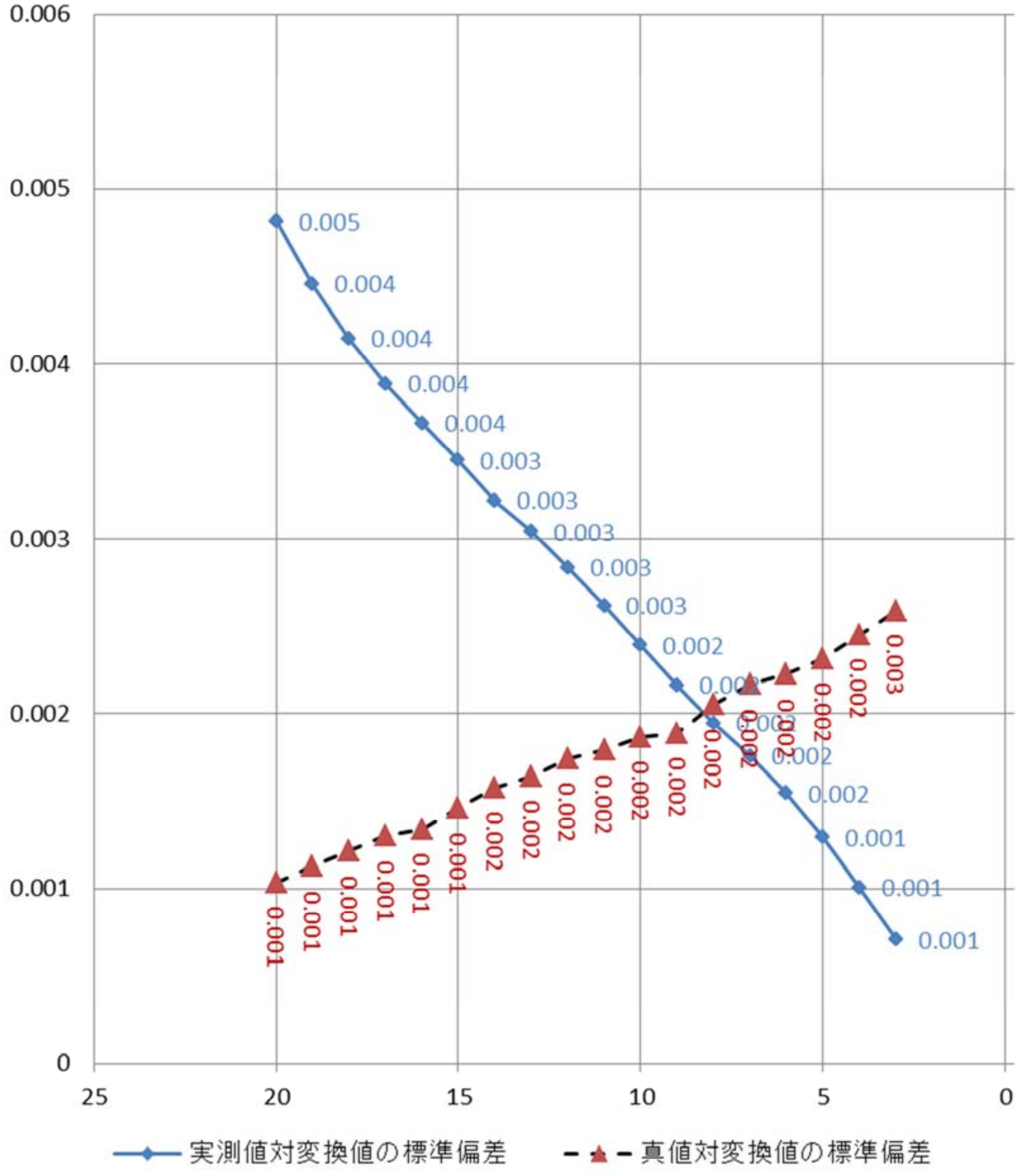


ここまでの説明で気がつかれたと思いますが，与えた標準偏差ベースの誤差もしくは母集団の標準偏差が判っていれば抜き取ったデータの標準偏差がその値以下でかつ最も近いデータの組合せの状態で最小二乗法座標変換すれば復元精度が高いことが判ることになりますが実務ではこの与えた標準偏差ベースの誤差，母集団の標準偏差は判りませんので出来ません。

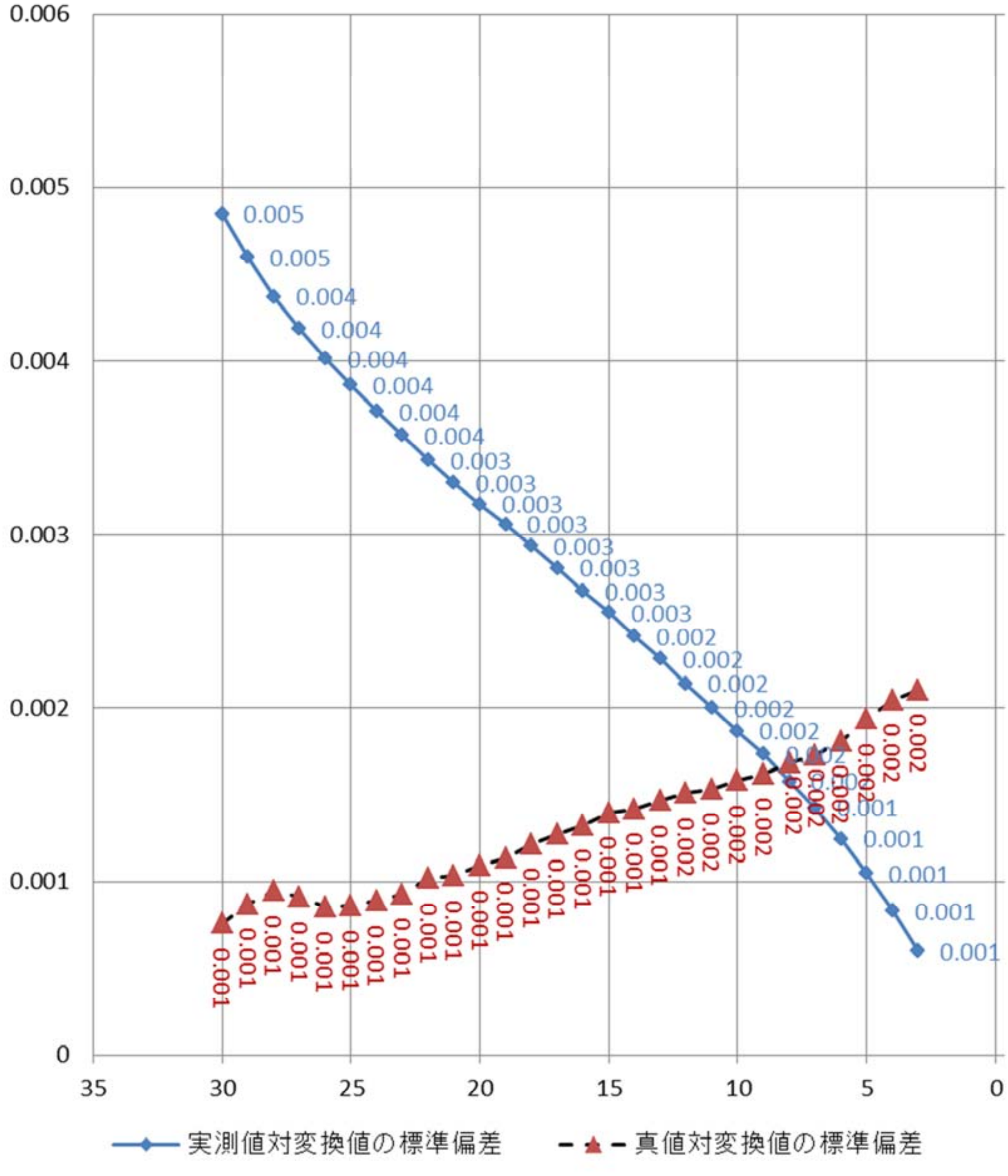
与えてある誤差は標準偏差ベースで0．0049，準拠点数10個～3個，30回の平均値



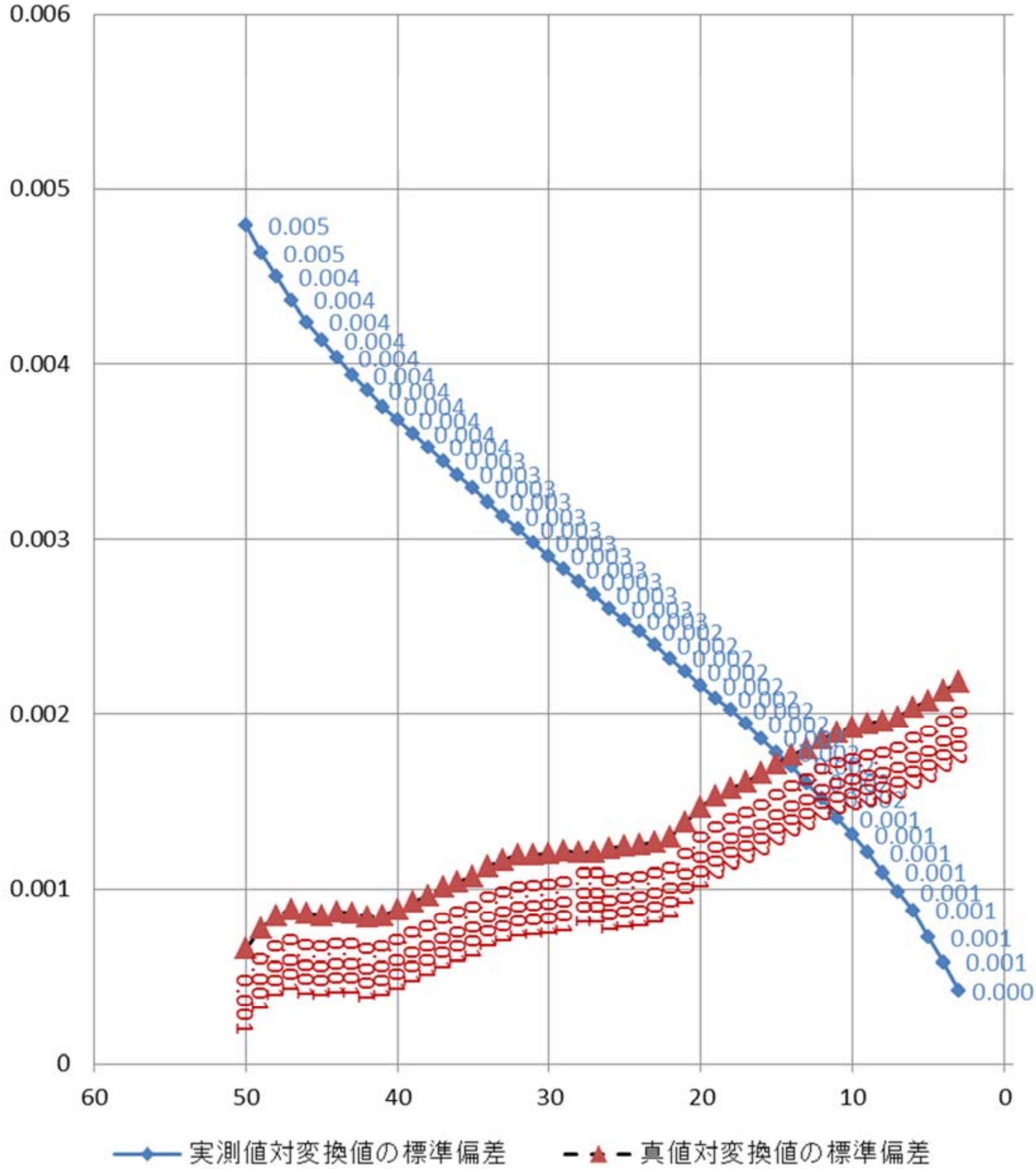
与えてある誤差は標準偏差ベースで0．0049，準拠点数20個〜3個，30回の平均値



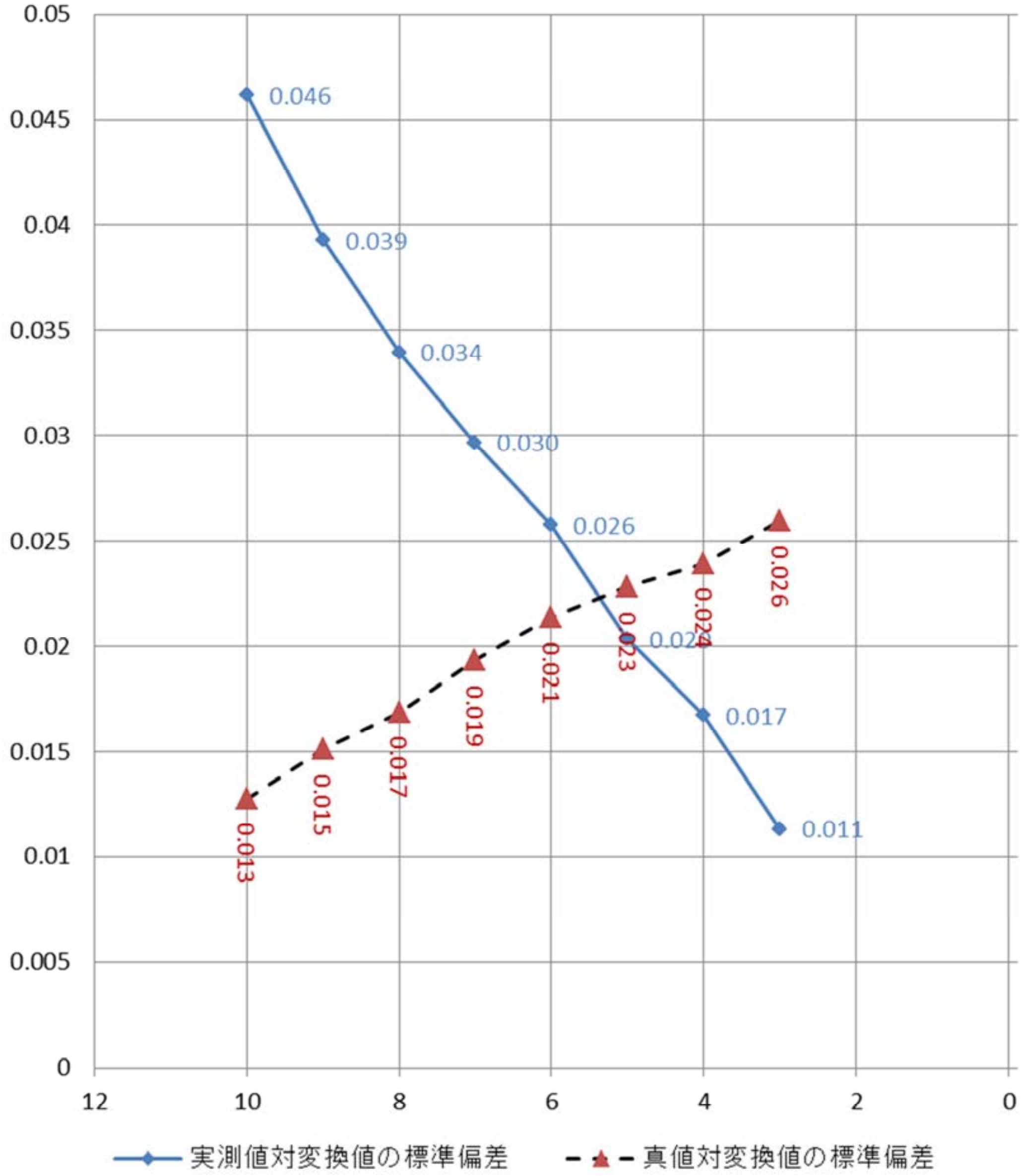
与えてある誤差は標準偏差ベースで0．0049，準拠点数30個～3個，30回の平均値



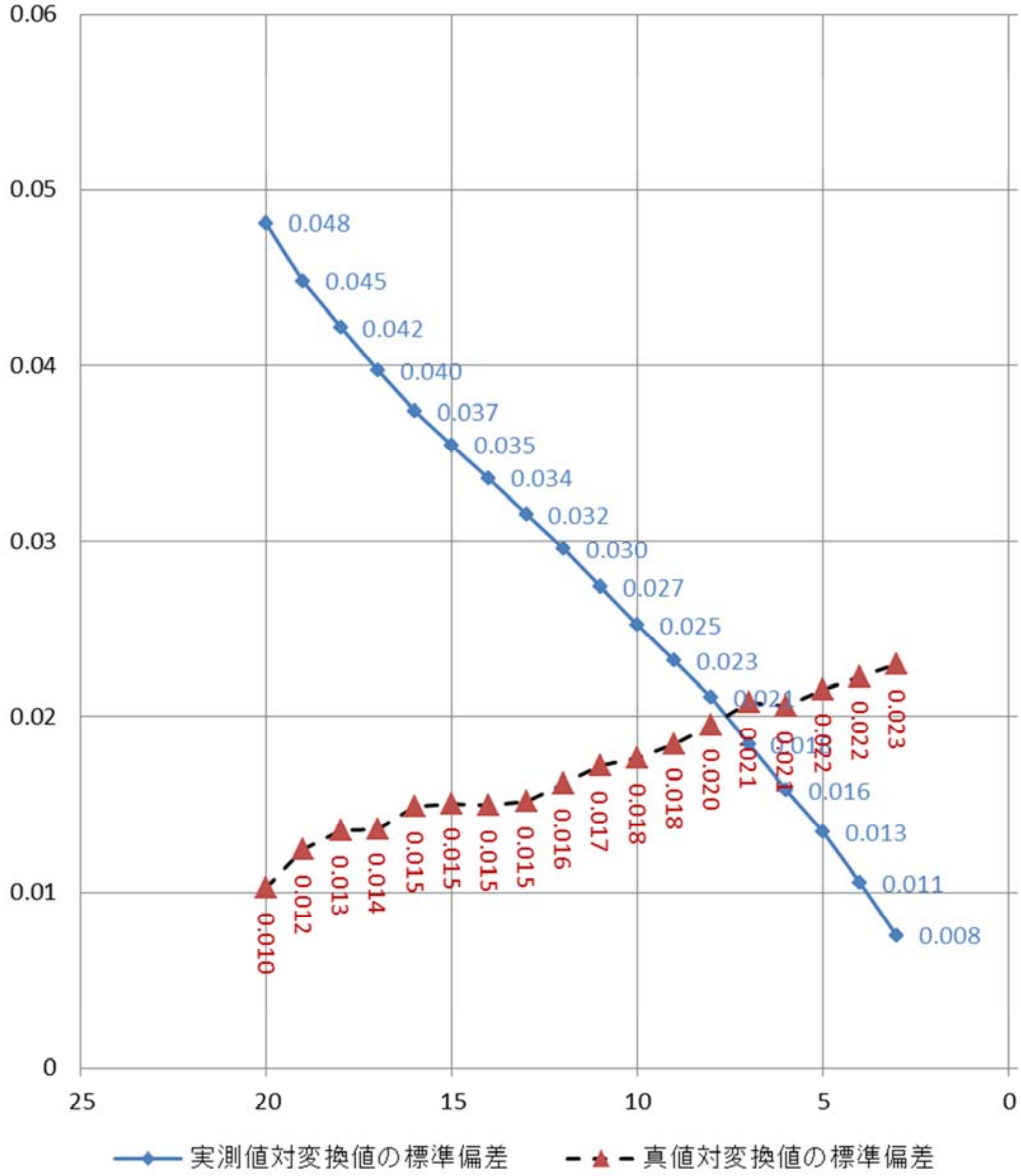
与えてある誤差は標準偏差ベースで0．0049，準拠点数50個～3個，30回の平均値



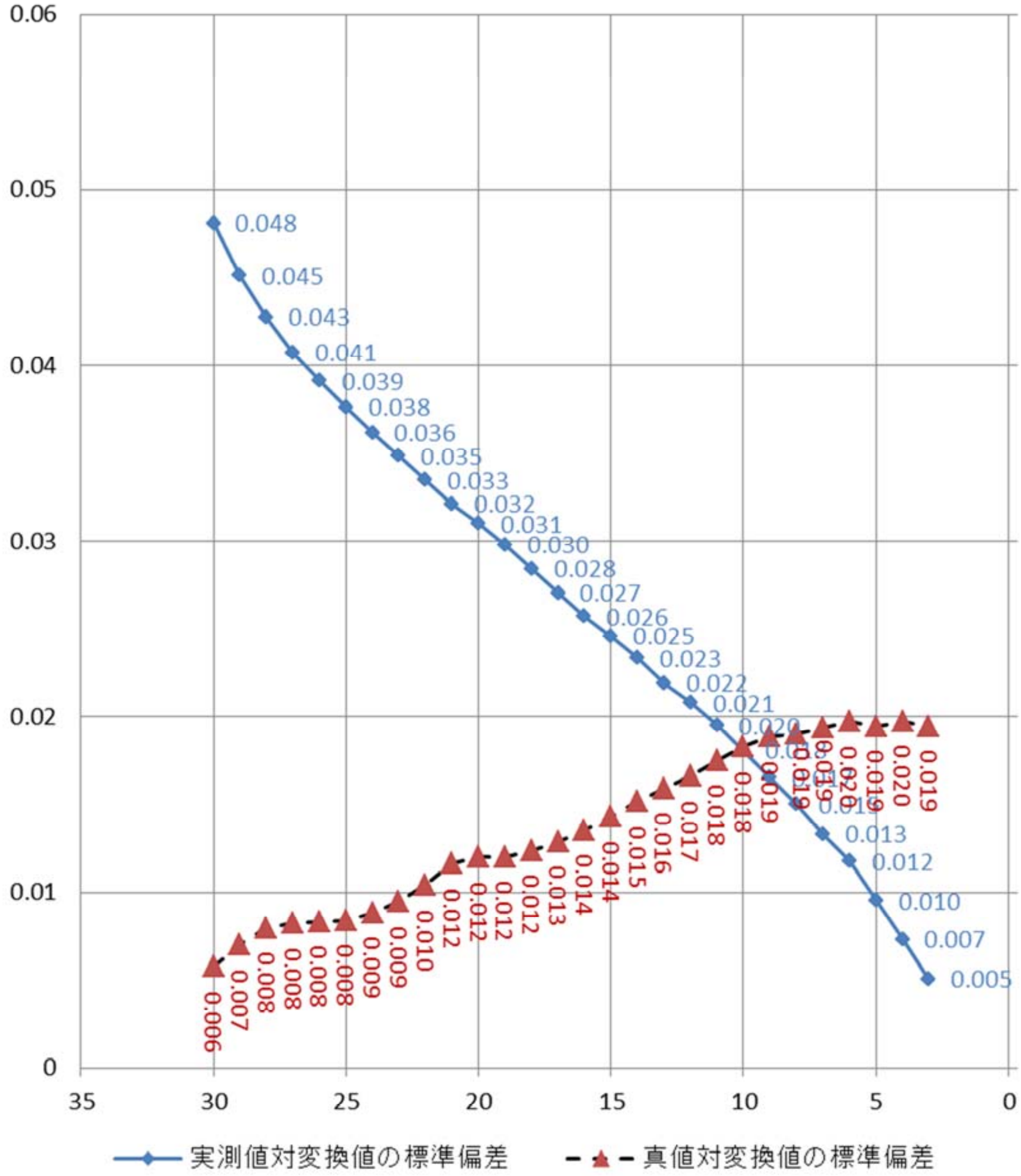
与えてある誤差は標準偏差ベースで0．0487，準拠点数10個～3個，30回の平均値



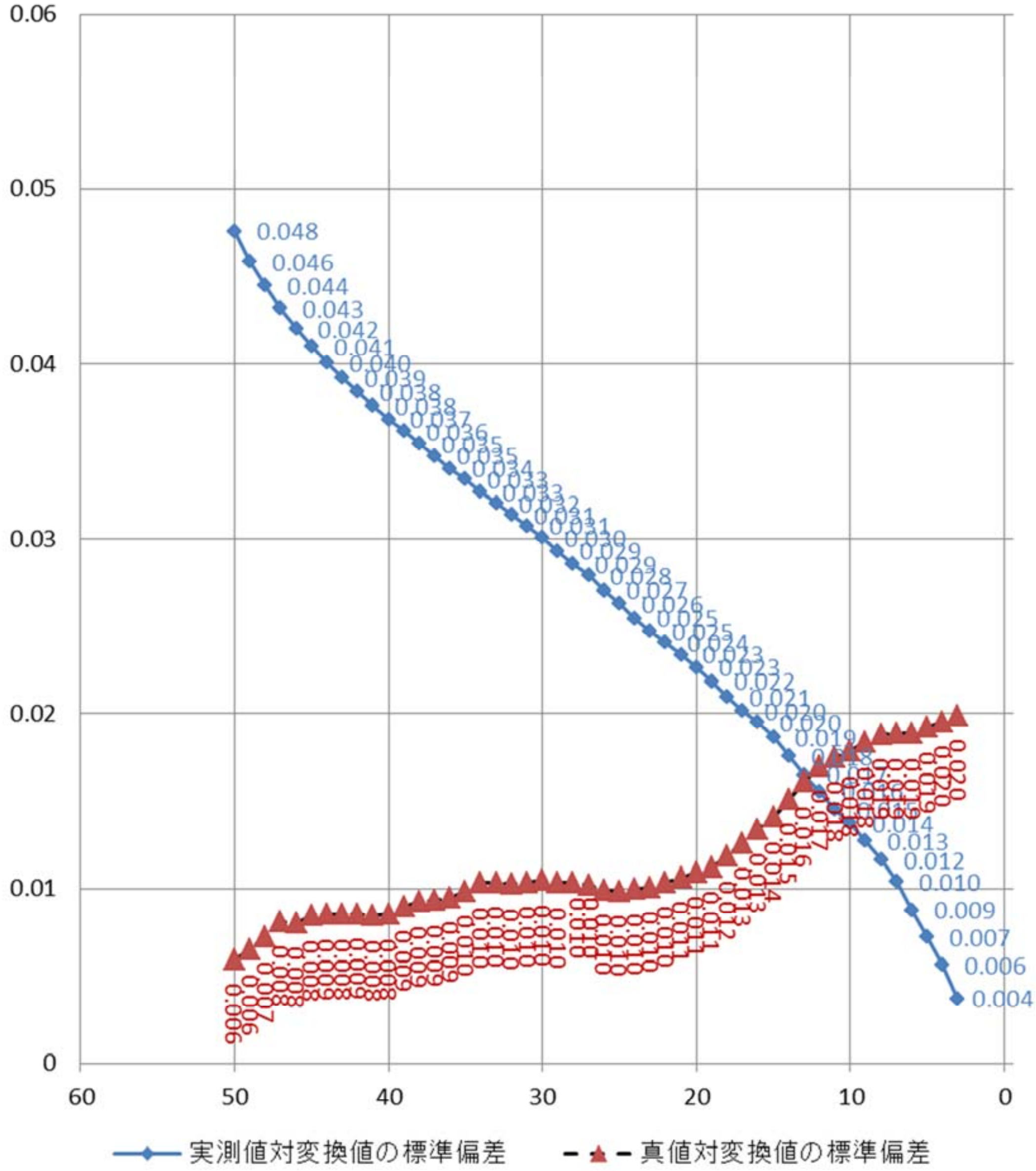
与えてある誤差は標準偏差ベースで0．0487，準拠点数20個～3個，30回の平均値



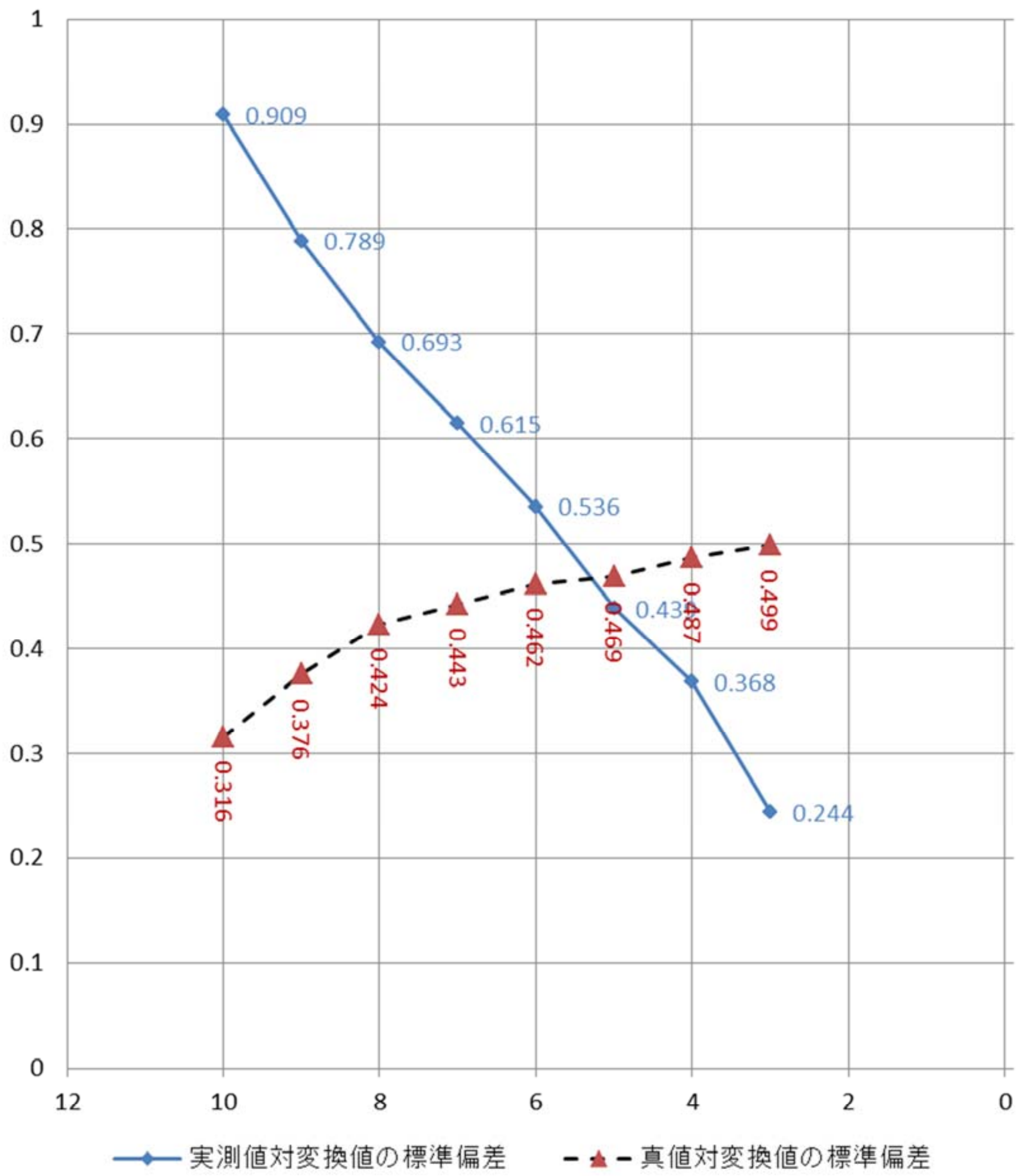
与えてある誤差は標準偏差ベースで0．0487，準拠点数30個～3個，30回の平均値



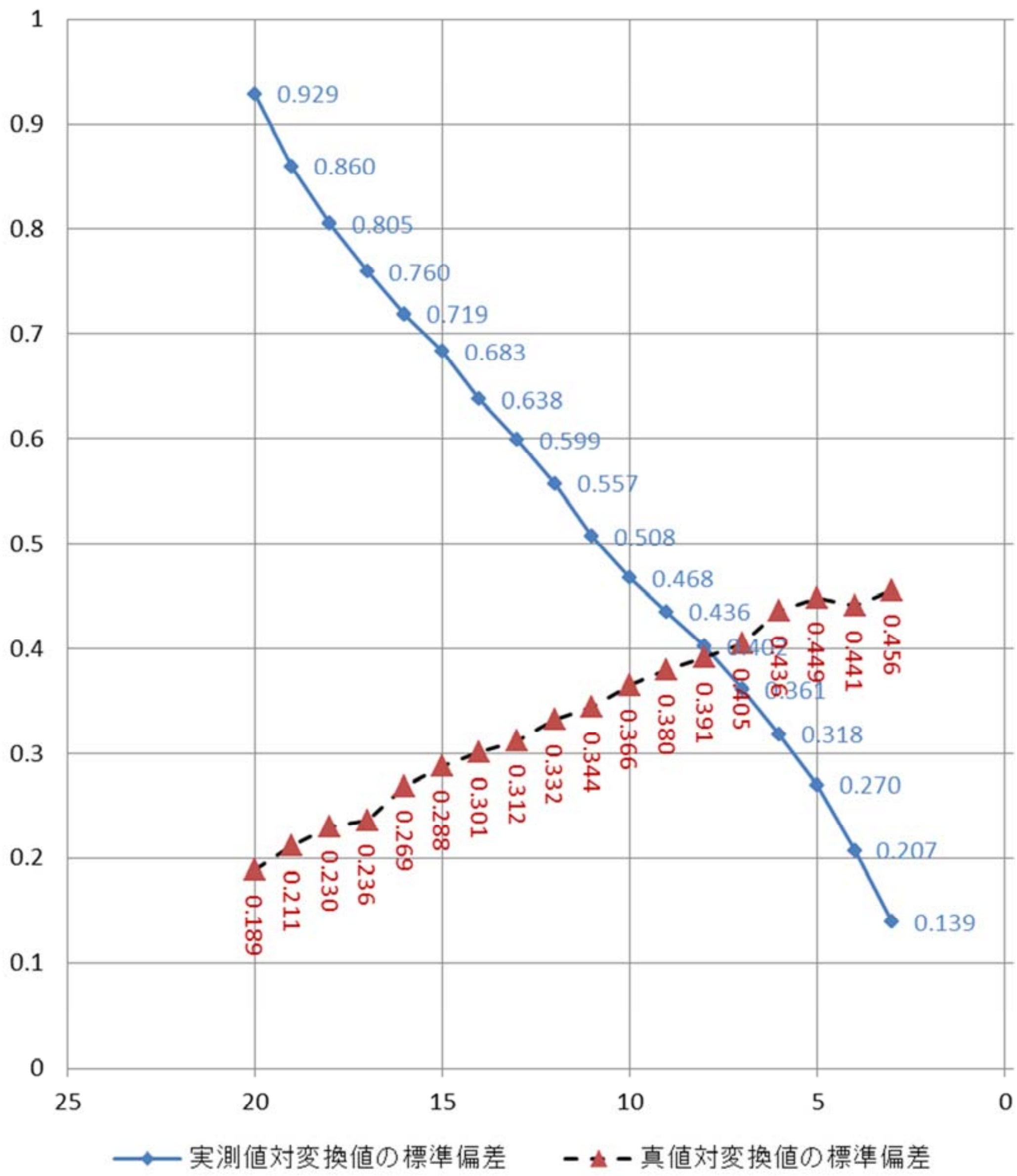
与えてある誤差は標準偏差ベースで0．0487，準拠点数50個～3個，30回の平均値



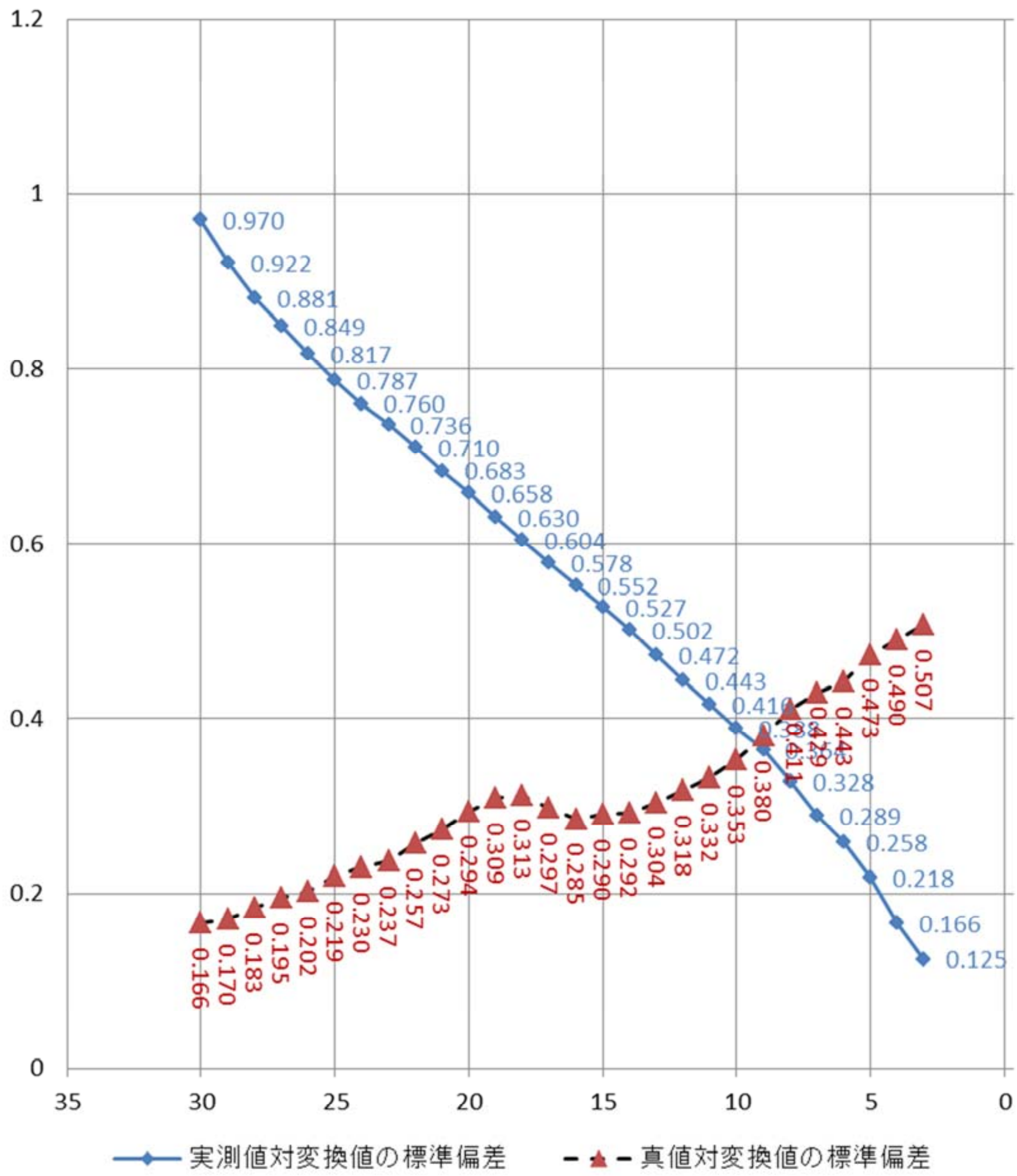
与えてある誤差は標準偏差ベースで0．9736，準拠点数10個～3個，30回の平均値



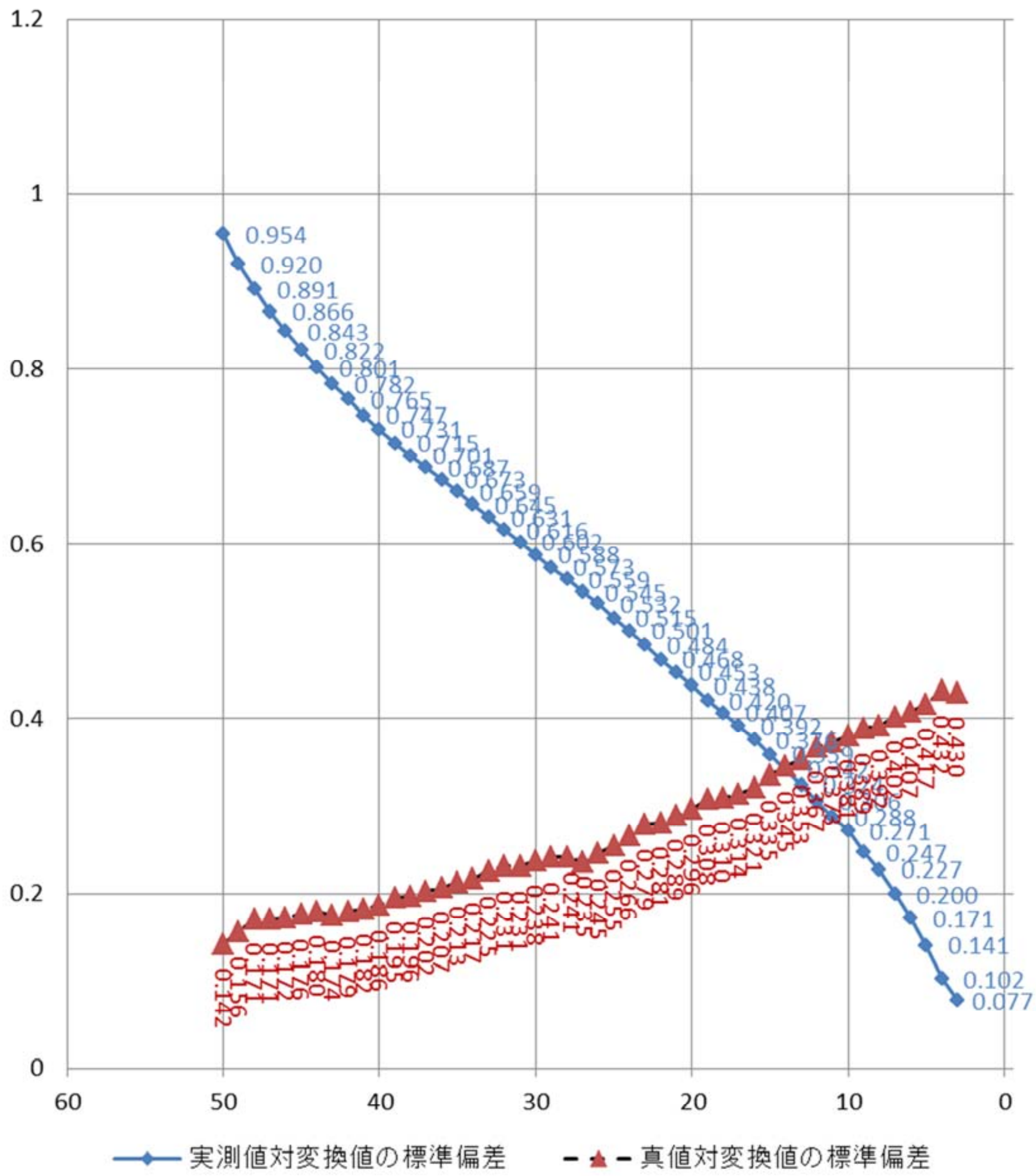
与えてある誤差は標準偏差ベースで0．9736，準拠点数20個～3個，30回の平均値



与えてある誤差は標準偏差ベースで0．9736，準拠点数30個～3個，30回の平均値



与えてある誤差は標準偏差ベースで0．9736，準拠点数50個～3個，30回の平均値



以上，計算結果をそのまま示しました，ある程度のバラツキを持ちながらも傾向はしっかりと出ていると言う点に注目してください。

実務では実測値と図面値の変換値の差（較差）の標準偏差を計算できますが本来の筆界との復元精度（真値と変換値の差（較差）の標準偏差）は知ることが出来ません。

準拠点選択へ

実測されたデータの中に正規分布に適合していないデータが含まれていなければ，母集団の精度，復元精度が推測出来ます，実際のデータには正規分布に適合していないデータが含まれている可能性がありますので確認して，あればそれを除くことが求められます。この方法「準拠点選択」について次に考えてみます。