

Title : 予測バンド95%(PB95%)についての一考

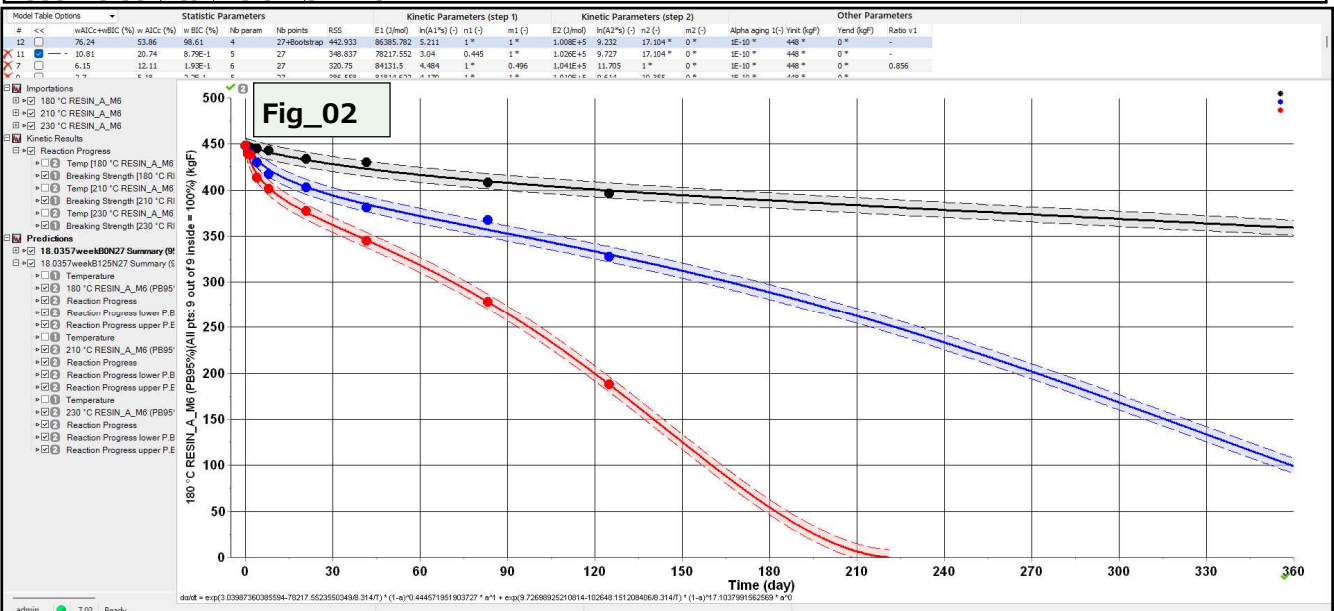
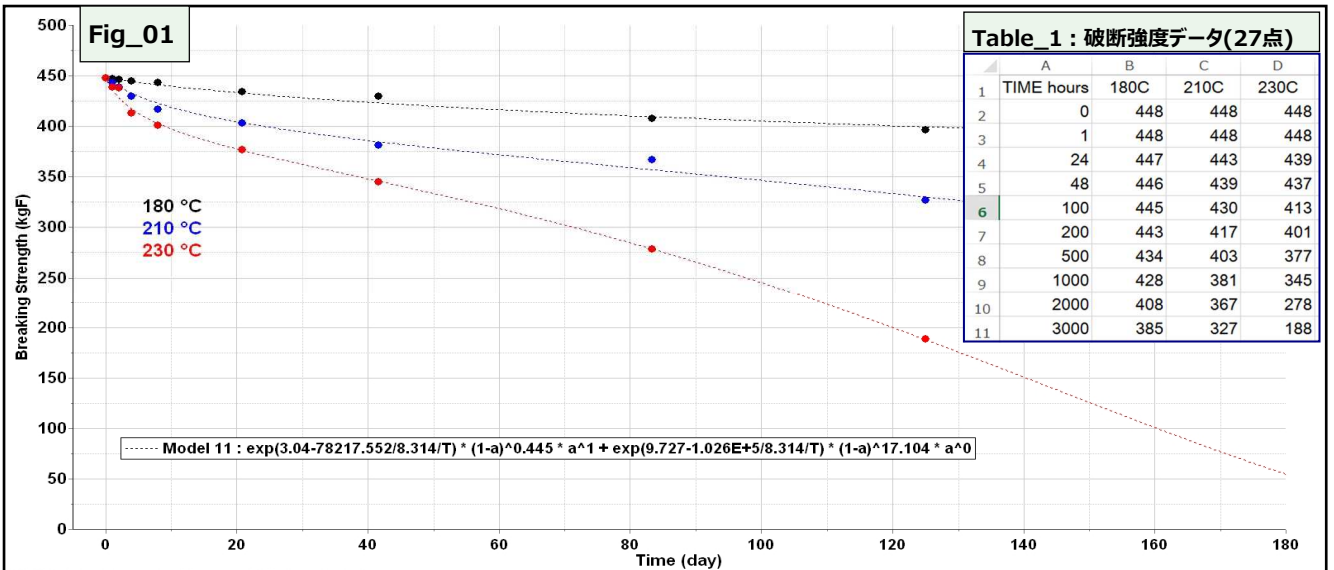


破断強度を測定する引張試験機

加速試験データの代表的なものとして、硬化性樹脂の破断強度から実使用条件下での長期寿命予測があります。得られた反応モデルと活性化エネルギーを用い、例えば使用温度80℃や60℃での破断強度が初期値の一定割合(例:1/10)まで低下するまでの時間を推定する。下図・Fig_01は等温条件180℃,210℃,230℃の引張試験の破断強度データ(27点)を使って、モデル式を求めた例です。実測データと回帰式の残差平均和RSSやAIC-BICの基準値のスコアからモデル式No.11が選択されました。反応モデル式は2stepの反応式です。活性化エネルギーは1step目が78.2kJ/mol、2step目が102.6kJ/molになっています。破断強度の最小値を0と仮定して得られたモデル式は実測データとのRSS(残差平方和)は348.83です。

この予測されたモデルの精度は残差解析(Residual Analysis)をクリックして得られた結果をFig_02に示します。この結果は残差の分散をもとにしたRSE(標準残差)を示し、“得られたモデル自身では説明できない典型的なずれの大きさ”を示します。言い換えるとRSE(標準偏差)はモデル式の当てはまり誤差(残差分散)のみを示しています。実測データ点が回帰曲線のまわりで、典型的にどの程度散らばっているか？ すなわち、誤差の“平均的な大きさ”を示します。しかし、新規の実測データがRSEの曲線上にプロットされることを意味するわけではありません。

#	<<	wAICc+wBIC (%)	w AICc (%)	w BIC (%)	Nb param	Nb points	RSS	E1 (J/mol)	ln(A1%) (-)	n1 (-)	m1 (-)	E2 (J/mol)	ln(A2%) (-)	n2 (-)	m2 (-)	Alpha aging 1(-)	Yinit (kgf)	Yend (kgf)
11	<input checked="" type="checkbox"/>	100	100	100	5	27	348.837	78217.552	3.04	0.445	1*	1.026E+5	9.727	17.104*	0*	1E-10*	448*	0*
2	<input type="checkbox"/>	OE+0	OE+0	OE+0	6	27	344.266	75426.045	3.015	1*	1.37	1.024E+5	9.641	16.266	0*	1E-10*	448*	0*



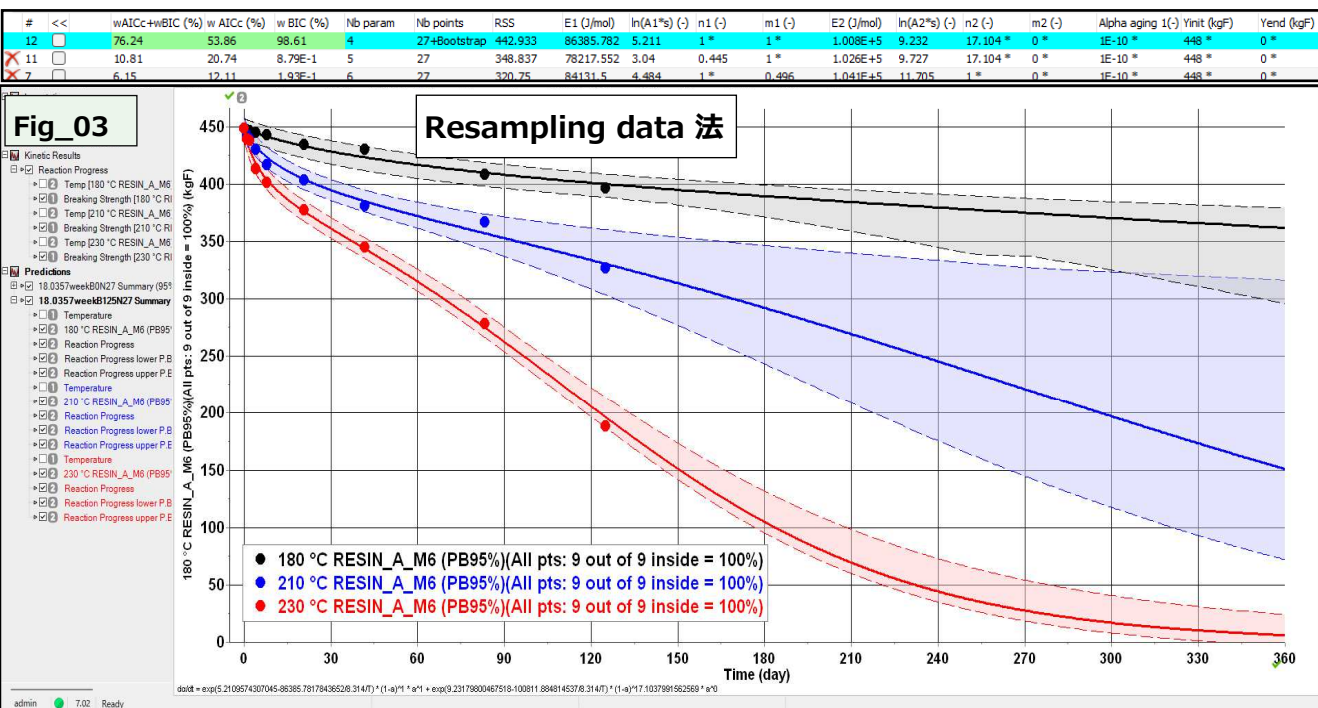
Technical Note テクニカルノート

No.TKsd-03/1 2025-10-03

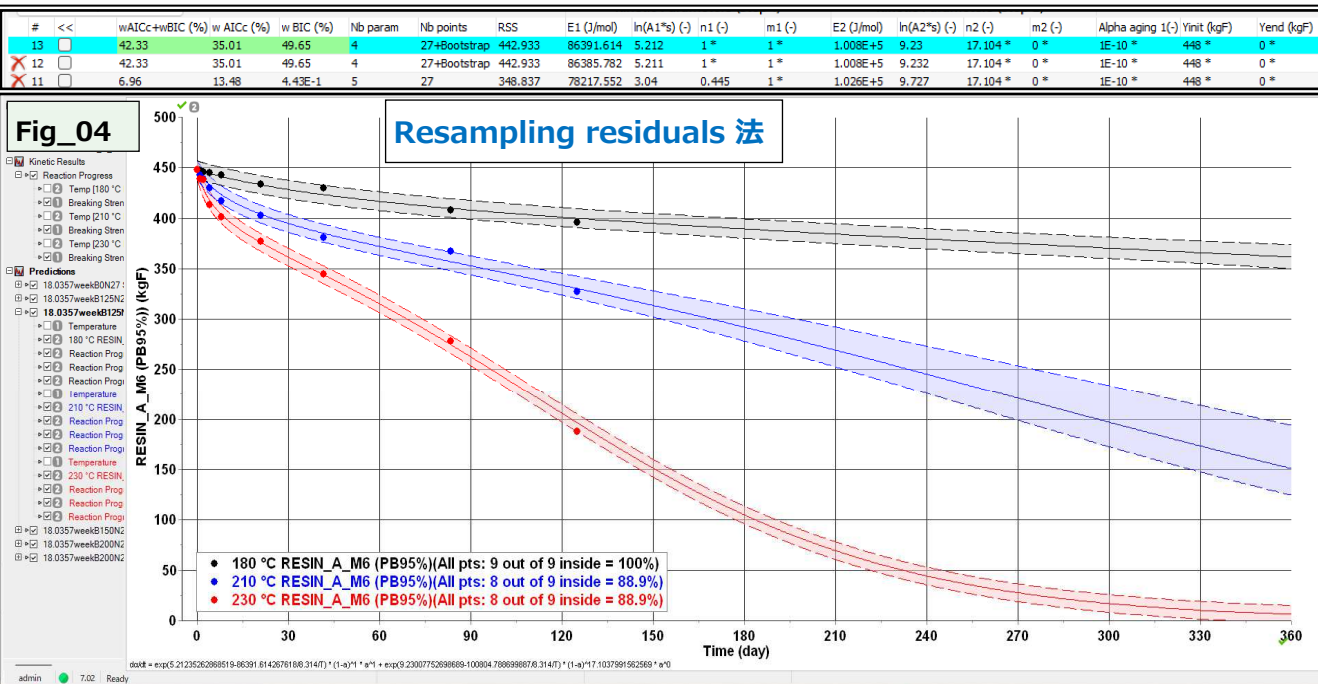
Title : 予測バンド95%(PB95%)についての一考

モデル式自体の誤差は得られても、新規に得られたデータがモデル式から予測される曲線上のどの位置にプロットされるのか？という課題があります。TKsd_V7では加速試験データの限られたデータ点数からモデルを求め、得られたモデル式から実際の保存期間を予測評価できるようにブートストラップ(Bootstrap)による予測バンドの計算機能が強化されました。

Bootstrapによる予測バンド解析はFig_03に示す①Resampling dataによる方法とFig_04に示す②Resampling residualsによる方法の2種類があります。



Bootstrap (Resampling data)計算をすると、Fig_03の上段のスコアが示すようにFig_02のMSE(標準残差)を求めたモデル式11からモデル式12に変わり、AIC-BICのスコア値や反応次数などがMSEを算出した時点のモデル式とは少なからず変化します。同様にBootstrap (Resampling residuals) 計算すると、Fig_04の上段のスコアが示すようにFig_03で求めたモデル式12からモデル式13に変化しますが、モデル式の変化はありません。その他の違いとして、Resampling dataでは実測データのすべてが予測バンド95%内に含まれたのに対し、Resampling residualsでは実測データのうち210℃と230℃の各1点がPB95%内に含まれなかったことが分かります。前者と後者のどちらのBootstrap計算手法を使うのが適切なのか？については次ページに続きます。



2つの計算方法のどちらを選ぶべきでしょうか？

●誤差の非正規性や異方分散、モデル・ミスペックの懸念がある場合

このような場合は **Resampling data** の方が非パラメトリックに全体のばらつきを捉え、過小評価リスクを抑えられます。
堅牢な予測バンドという意味では**Resampling data**法が過小評価リスクが抑制されていることになります。

モデルが正しく指定され、残差が“independent and identically distributed” の場合 **残差が独立同一分布の場合**

● independent(独立)とは、各データ点が互いに影響を与えずに得られている。

● identically distributed(同一分布)とは、すべてのデータ点が同じ確率分布から生成されている。

このような場合は **Resampling residual** のほうが効率的で、真の誤差構造に即した予測バンドを得やすくなります。

予測バンドで表示されたグラフが何を意味するか？を言い換えると新規サンプルの値がどこまで信頼して予測できるかを示す区間を示します。
またRSEによるバンド幅と予測バンド(PB95%)を比較した場合、その予測バンド幅の方が広がります。

Bootstrapによる予測バンド算出のメリット・デメリットについて

メリット

○ 厳密な分布仮定が不要なノンパラメトリック手法である。ノンパラメトリック手法とは、データの母集団分布をあらかじめ特定の確率分布(正規分布など)やパラメータ(平均や分散など)に仮定せずに解析を行う方法を指します。

○ 元データ数が小さくても適用可能で、小サンプルの事例に強い。

○ 非線形モデルや異分散、外れ値を含む複雑なデータ構造にも柔軟に対応できる。

○ 各説明変数の点(xごと)で、応答の分布を直接推定できる。

パラメトリック手法では難しいモデル誤差の非対称性やバイアスを補正する拡張(BCa法など)も利用できる。

デメリット

● 再標本化回数に比例して計算コストが上昇し、大規模データや高次元モデルでは処理時間が問題になる。

● 標本データが母集団を代表していない場合、予測バンド幅の信頼性が低下しやすい。

● 再標本化の方法や回数設定が結果に影響し、選択基準の明確化が必要になる。

● そのままではバイアスが残る場合があり、BCa法などによる補正を検討する必要がある。

時系列・空間データには標準的のブートストラップでは対応しにくく、ブロックブートストラップなど特殊手法が求められる。

残差分析で得られる「残差の標準誤差(Residual Standard Error; RSE)」と、95%予測バンドが示す内容は目的も計算式も異なります。以下で両者を対比しながら解説します。

残差の標準誤差(Residual Standard Error; RSE)とは

回帰モデルでは、実際に測定した値とモデルが予測した値に必ずズレが生じます。このズレを「残差」と呼びます。

たとえば時間と温度の関係を予測した際に、実際の温度がモデルの予測値より高かったり低かったりする差が残差です。

残差の標準誤差は、残差がどれくらい散らばっているかをまとめて示す指標です。

言い換えると、モデルが説明しきれない誤差の「だいたいこれくらい」の大きさをひとつの数字で表しています。

残差のばらつきが小さければ、モデルがデータをよく捉えていると判断できます。

この指標はモデルの当てはまりの良さを評価する目安になり、回帰係数の検定や信頼区間の計算など、統計的な判断を行う基礎データとしても利用されます。残差の標準誤差が大きい場合は、モデルの見直しや説明変数の追加を検討するきっかけになります。

予測バンド95%(PB95%)とは

ある条件(例えば特定の温度や時間)で観測を行ったとき、新しく得られる値がどのくらいの範囲に収まりやすいかを示す帯です。

以下のように理解するとわかりやすいでしょう。

* 回帰線は「過去のデータの平均的な傾向」を示します。

* 95%予測バンドはその回帰線の両側に描かれるバンド幅で、「これから観測する値が95%の確率でここに入りますよ」という目安です。

また予測バンドは、残差の標準誤差の幅よりもバンド幅が広がります。

バンド幅が決まるポイントについて

* データのばらつきの大きさは、データがバラバラに散らばっているほどバンド幅は広がります。

* データのサンプル数は実測データが少ないと不確実性が高くなるため、バンド幅は広がります。

* 回帰線からの距離 回帰線から遠い条件ほど予測の信頼性が下がるため、バンド幅が広がります。

バンド幅は何に使えるのか？

* 未来の観測値の範囲を見積もる。

* 観測値がこのバンド幅の外に出たときに、予測モデルや測定に何か問題がないかを確認する。

* このように、95%予測バンドは「モデルがどれだけ安心して使えるか」を直感的に示してくれる道具です。

今まで当社におけるTKsdの解析作業は、DSCデータの等温条件の反応率曲線から数100点の間引きデータを作成してモデル式を探索する使い方であったため、回帰計算によるモデル式を得ることが主目的で単純な課題でした。

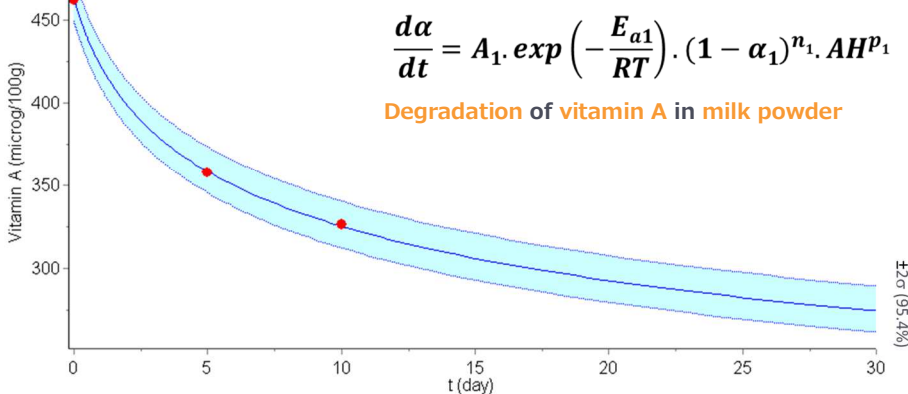
これがバイオ医薬品の保存期間評価となるとバイオ医薬品の加速試験データから得られた**モデル式、予測バンド幅および検証データ**を使って、解析対象となる新規の実測データから得た情報から劣化状態と保存期間の関係について、高度な判断が求められることになります。

検証データについて 次ページに具体例を紹介します。

Title : 予測バンド95%(PB95%)についての一考

Fig_05

Iso (60°C,RH60%) PB95% Bootstrap (Resampling data)



この解析事例は2023年に開催された Swiss Society for Thermal Analysis and Calorimetry (STK) で発表された内容の一部を転記しています。食品(ミルク)中のビタミンAの熱分解反応を25°C・60%RH, 37°C・60%RH, 60°C・60%RHの等温条件で加速試験を行い、TKsdソフトウェアにより劣化反応のモデル式を探索しています。

実測データ数は60°C等温条件 3点 37°C等温条件5点, 25°C等温条件4点合計で12点です。その他25°C等温条件の検証用データが3点 による解析結果です。

ミルク100 g 中のビタミン A の重量は Fig_05の実測データから 60°C等温条件で10日後に330μg, Fig_06の実測データから 37°C等温条件で150日後に390μg, Fig_07の実測データから 25°C等温条件で9か月後に445μg, の劣化によるビタミンAの減量変化が検出されています。

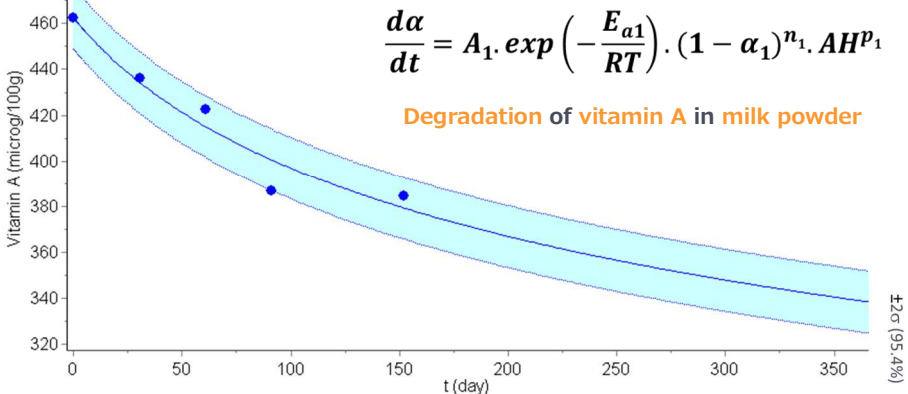
得られたモデル式は
A1 : 30.693
Ea1 : 128.4J/mol
n1 : 7
が得られました。

温度水準が3条件と実測データ合計 12点からAIC-BICのスコア評価は低いですが、モデルを特定できています。

Fig_07には得られたモデルが妥当性があるか否かを判断するための検証データ 3点が○で表示されています。検証用のデータがBootstrapの計算が Resampling data法で得られた PB95%・予測バンド(95.4%)のバンド幅に収まっていることが確認できます。検証データにより、150日以降の保存期間であっても予測の判定が可能であることが担保されます。検証データより先の期間の予測精度を担保するには検証用データが必須であることを意味します。

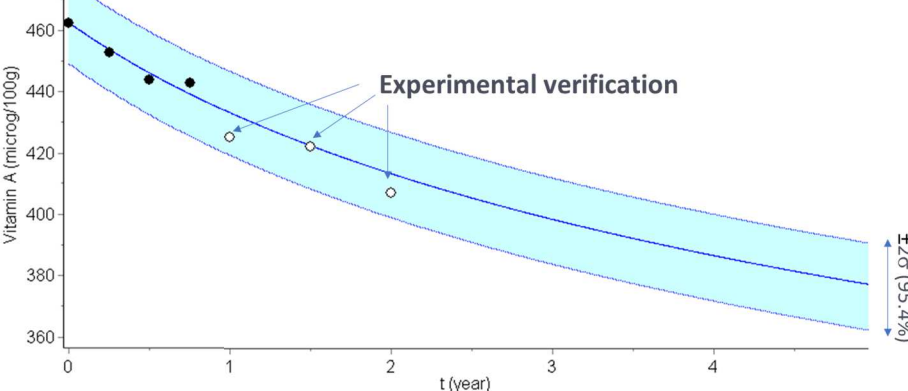
Fig_06

Iso (37°C,RH60%) PB95% Bootstrap (Resampling data)



Fig_07

Iso (25°C,RH60%) PB95% Bootstrap (Resampling data)



Fig_07の事例では1年~5年後の予測幅の広がり方が軽微ですが、Fig_03の210°Cの例のように時間経過に従って予測バンド幅が増大することもあります。予測幅が広がった場合、予測することに妥当性があるのか？を再検討する必要があることもあり得ます。



もし、お客さまで加速試験データセットを提供された場合、
劣化モデル式がどのようになるか？
実測データの精度がどれくらいか？
予測バンド95%のバンド幅など
について解析結果を報告します。

Palmetrics

〒357-0038 埼玉県飯能市仲町127-20 リバーサイド飯能 202
株式会社パルメトリクス はんのう研究室
電話 042-978-8655 FAX 042-978-8664