T 法における欠測データの活用に関する研究

Study of T-Method for Missing Data

有限会社 増田技術事務所　増田 雪也

1. はじめに

T 法[1]は誰でも簡単に重回帰分析ができる手法である。数式が単純なため、MS-Excel でも容易に解析を行うことができ、様々な分野で活用され始めている。

あまり知られてはいないが、T 法には従来の重回帰分析にはないメリットがある。それは、「欠測データが存在していても解析できる」というメリットである。しかし、欠測データ数が、T 法における重み付けの SN 比 η に悪影響を与える可能性があると考えられる。

そこで本研究では、重み付けの SN 比 η を求める際、エネルギー比型 SN 比[2]を用いて解析する方法を検討した。その結果、従来の SN 比との差は非常に小さく、推定精度に与える影響は軽微であった。

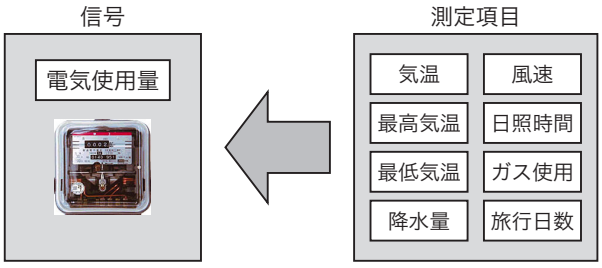


図１　我が家の電気使用量の解析（信号と測定項目）

2. 欠測データの活用方法

2.1 事例「我が家の電気使用量の解析」

T 法における欠測データの活用方法を説明する。その事例として、「我が家の電気使用量の解析」を取りあげる。この事例は、図 1 に示すように、信号として「電気使用量（※月毎に日数が異なるため、毎月の電気使用量から 1 日当たりの平均使用量を求めたもの）」、測定項目として、月毎の「平均気温、平均最高気温、平均最低気温、降水量、平均風速、日照量、ガス使用量、旅行で留守にした日数」を設定した。気象データは、気象庁の Web サイトから得た。

図 2 に 2004 年 9 月から 2011 年 6 月までの電気使用量を示す。冬季になると電気使用量が高くなるのは、暖房としてエアコンを使用しているためである。我が家は、長野県の高原にあるため、夏期にエアコンを使用することはない。

この全データの内、2004 年 9 月から 2005 年 8 月までの 12 ヶ月分のデータを、既知データとして用い、残りの 2005 年 9 月から 2011 年 6 月までの 70 ヶ月分のデータを、未知データとして推定することとした。

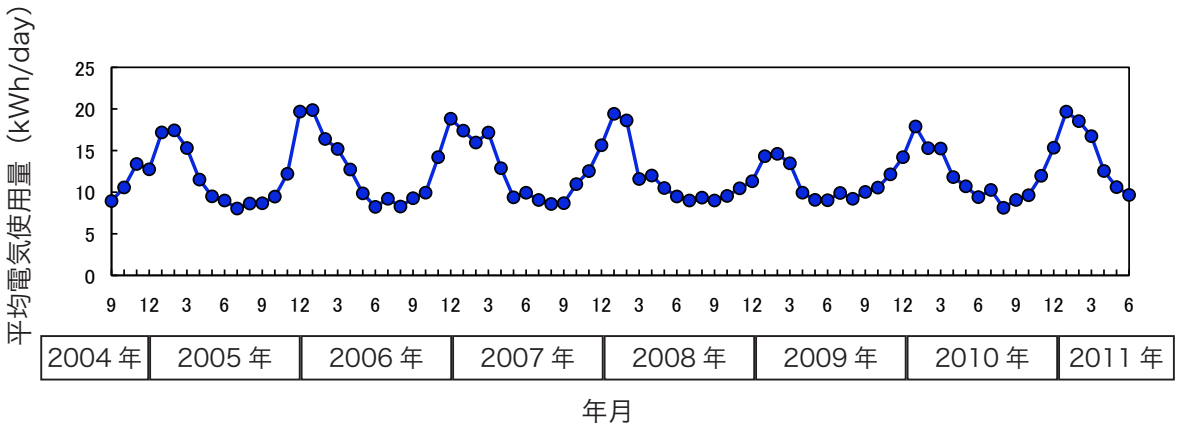
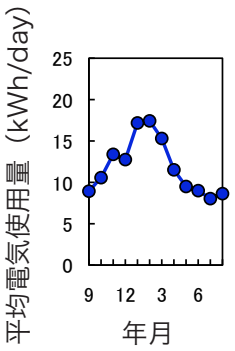


図２　2004 年から 2011 年までの電気使用量

Setsuya Masuda (info2qe@abox3.so-net.ne.jp)
Masuda Engineering Consultant Office,Inc.



	真値	測定項目							
	M	k1	k2	k3	k4	k5	k6	k7	k8
年月	電気使用量	ガス	降水量	気温	最高気温	最低気温	風速	日照量	旅行日数
2004.9	8.93	8.5	5.1	18.8	24.2	14.9	1.6	4.7	4
2004.10	10.57	9.2	9.3	10.9	16.3	6.5	1.6	4.5	3
2004.11	13.39	12.3	1.8	6.8	13.0	1.8	1.6	6.0	3
2004.12	12.75	6.6	2.2	1.1	7.2	-3.8	1.8	6.3	18
2005.1	17.18	13.2	1.3	-3.7	1.5	-9.0	1.8	5.4	5
2005.2	17.43	14.2	1.1	-2.7	2.1	-7.5	2.0	5.6	0
2005.3	15.30	15	2.8	0.9	6.5	-4.7	2.3	6.1	2
2005.4	11.52	15	1.5	8.8	16.1	1.7	2.2	7.3	0
2005.5	9.50	9	1.8	12.5	19.6	6.3	2.0	6.9	7
2005.6	9.00	10.6	3.9	18.7	24.7	14.0	1.5	4.9	0
2005.7	8.03	8.2	5.4	21.1	26.5	16.7	1.4	4.4	7
2005.8	8.64	8.5	2.0	21.9	27.7	18.1	1.3	5.0	2

図３　T 法の推定式を作るための既知データ

図３に、T 法の推定式を作るための既知データ（12 ヶ月分）を示す。

2.2 重回帰分析における欠測データへの対応

図４に示すように、従来の重回帰分析では、欠測値が存在する場合は、欠測値を含む行（つまり、当該月のデータ）を丸ごと削除して解析する。あるいは、列方向（つまり測定項目）の欠測値が多い場合は、その列の測定項目を丸ごと削除して解析する。

図５に示すのは、極端な事例であるが、斜めに欠測値が存在する場合や、市松模様のような場合は、従来の重回帰分析では解析が難しい。

削除

	真値	測定項目							
	M	k1	k2	k3	k4		k6	k7	k8
年月	電気使用量	ガス	降水量	気温	最高気温		風速	日照量	旅行日数
2004.9	8.93	8.5	5.1	18.8	24.2		1.6	4.7	4
欠測									
2004.11	13.39	12.3	1.8	6.8	13.0		1.6	6.0	3
欠測									
2005.1	17.18	13.2	1.3	-3.7	1.5		1.8	5.4	5
2005.2	17.43	14.2	1.1	-2.7	2.1	欠測	2.0	5.6	0
2005.3	15.30	15	2.8	0.9	6.5		2.3	6.1	2
2005.4	11.52	15	1.5	8.8	16.1		2.2	7.3	0
2005.5	9.50	9	1.8	12.5	19.6		2.0	6.9	7
2005.6	9.00	10.6	3.9	18.7	24.7		1.5	4.9	0
2005.7	8.03	8.2	5.4	21.1	26.5		1.4	4.4	7
2005.8	8.64	8.5	2.0	21.9	27.7		1.3	5.0	2

削除　削除

図４　重回帰分析における欠測データへの対応

2.3 T 法における欠測データへの対応

T 法の特徴は、「SN 比による重み付け」をしていることである。図６に、その SN 比の概念を示す。信号と項目の関係に対してゼロ点を通る直線を設定し、この直線からのズレを数値化したものが『ゼロ点比例式の SN 比』である。T 法では、この SN 比の値を「重み付け」として用いている。SN 比の詳細については、参考文献[1)]を参照されたい。

図７の３つのグラフは、信号と測定項目の関係における代表的な３つのパターンである。このように、バラツキ具合を SN 比によって重み付けしている。

斜めに欠測

	真値	測定項目							
	M	k1	k2	k3	k4	k5	k6	k7	k8
年月	電気使用量	ガス	降水量	気温	最高気温	最低気温	風速	日照量	旅行日数
2004.9	8.93	8.5	5.1	18.8	24.2	14.9	1.6	4.7	4
2004.10	10.57	9.2	9.3	10.9	16.3	6.5	1.6	4.5	3
2004.11	13.39	12.3	1.8	6.8	13.0	1.8	1.6	6.0	3
2004.12	12.75	6.6	2.2	1.1	7.2	-3.8	1.8	6.3	18
2005.1	17.18	13.2	1.3	-3.7	1.5	-9.0	1.8	5.4	5
2005.2	17.43	14.2	1.1	-2.7	2.1	-7.5	2.0	5.6	0
2005.3	15.30	15	2.8	0.9	6.5	-4.7	2.3	6.1	2
2005.4	11.52	15	1.5	8.8	16.1	1.7	2.2	7.3	0
2005.5	9.50	9	1.8	12.5	19.6	6.3	2.0	6.9	7
2005.6	9.00	10.6	3.9	18.7	24.7	14.0	1.5	4.9	0
2005.7	8.03	8.2	5.4	21.1	26.5	16.7	1.4	4.4	7
2005.8	8.64	8.5	2.0	21.9	27.7	18.1	1.3	5.0	2

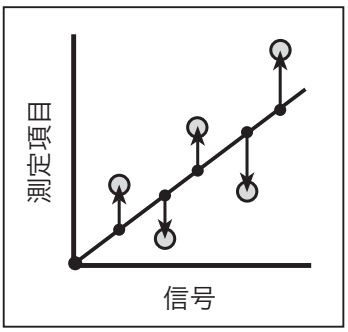
解析は難しい

市松模様で欠測

	真値	測定項目							
	M	k1	k2	k3	k4	k5	k6	k7	k8
年月	電気使用量	ガス	降水量	気温	最高気温	最低気温	風速	日照量	旅行日数
2004.9	8.93	8.5	5.1	18.8	24.2	14.9	1.6	4.7	4
2004.10	10.57	9.2	9.3	10.9	16.3	6.5	1.6	4.5	3
2004.11	13.39	12.3	1.8	6.8	13.0	1.8	1.6	6.0	3
2004.12	12.75	6.6	2.2	1.1	7.2	-3.8	1.8	6.3	18
2005.1	17.18	13.2	1.3	-3.7	1.5	-9.0	1.8	5.4	5
2005.2	17.43	14.2	1.1	-2.7	2.1	-7.5	2.0	5.6	0
2005.3	15.30	15	2.8	0.9	6.5	-4.7	2.3	6.1	2
2005.4	11.52	15	1.5	8.8	16.1	1.7	2.2	7.3	0
2005.5	9.50	9	1.8	12.5	19.6	6.3	2.0	6.9	7
2005.6	9.00	10.6	3.9	18.7	24.7	14.0	1.5	4.9	0
2005.7	8.03	8.2	5.4	21.1	26.5	16.7	1.4	4.4	7
2005.8	8.64	8.5	2.0	21.9	27.7	18.1	1.3	5.0	2

解析は難しい

図５　極端な事例



ゼロ点を通る直線からのズレ

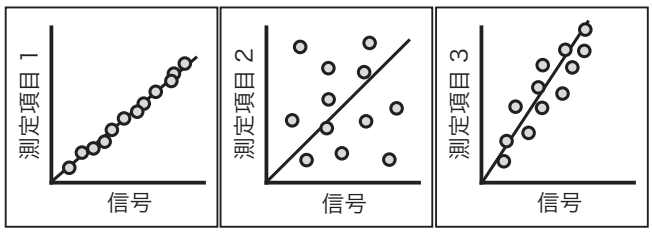
ゼロ点比例式のＳＮ比

【重み付け】

図６　【重み付け】として用いるＳＮ比の概念

図８の左のグラフは、欠測値が無い場合である。一方、右のグラフは、左のグラフから５つのデータを欠測させた場合である。このケースでは、例え５つの欠測が発生したとしても、全体的なバラツキの大きさは、ほぼ同程度であり、ゼロ点比例式で求められる重み付けの SN 比 η も、ほぼ同程度となる。データが欠測しても、データ数が少なくなるだけで、SN 比の算出に数学的な問題は無い。従ってＴ法では、欠測値が存在しても、重み付けの SN 比を求めることができ、推定値を得ることが可能となる。ただし、推定精度の良し悪しは、ケースバイケースである。

図９に示すように、３つの測定項目において、それぞれ欠測数が異なったとしても、重み付けの SN 比を求めることができ、各 SN 比 η と各傾き β から、推定値 $\hat{M}_i$ を求めることができる。なお、図９の数式は、一見複雑そうに見えるが、各測定項目の重み付けの SN 比で加重平均を求めているにすぎない。



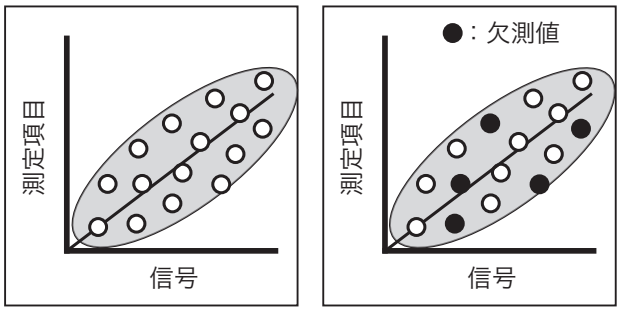
活用できるデータ　活用できないデータ　そこそこ活用できるデータ

重み付け：重 ＝ ＳＮ比：高

重み付け：軽 (0) ＝ ＳＮ比：低 (0)

重み付け：中 ＝ ＳＮ比：中

図７　Ｔ法の特徴【重み付け】



ゼロ点比例式の SN 比 η　　ゼロ点比例式の SN 比 η

データが欠測しても算出できる

図８　欠測データが存在する場合の重み付け

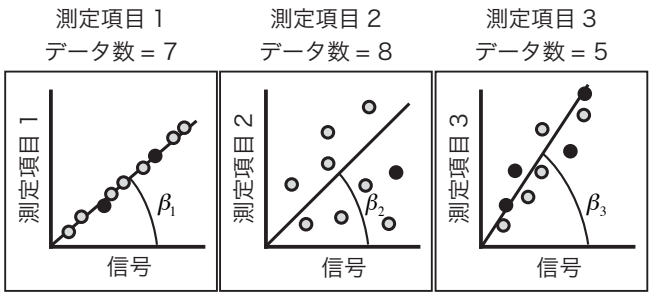
2.4 故意にデータを欠測させた場合の比較

図 10 に、欠測値の有無および欠測の程度の違いによる、総合推定の SN 比および相関係数の違いを示す。左は欠測が無い場合である。中央は故意に斜めに欠測させた場合、右は市松模様に欠測させた場合である。この「我が家の電気使用量の解析」事例では、欠測値の有無および程度の違いによらず、総合推定の SN 比および相関係数は、ほぼ同じ値となった。

3. 欠測データを活用する上での問題点

3.1 欠測データ数の影響

図９に示すような場合、データ数は測定項目毎に異なっている。式 (1) に T 法の重み付けの SN 比を示す。式 (1) 中の Ve は、式 (2) によって求められる。この Ve は、データ数 l の影響を受ける。従って、測



SN 比 $\eta_1 = 1.7$　　SN 比 $\eta_2 = 0.05$　　SN 比 $\eta_3 = 1.1$

$$\text{推定値 } \hat{M}_i = \frac{\eta_1 \times \frac{x_{i1}}{\beta_1} + \eta_2 \times \frac{x_{i2}}{\beta_2} + \eta_3 \times \frac{x_{i3}}{\beta_3}}{\eta_1 + \eta_2 + \eta_3}$$

データが欠測しても、推定値は算出できる

図９　欠測データが存在する場合の推定値の算出

欠測無し

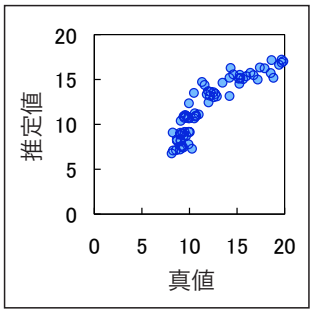
年月	真値 M 電気使用量	測定項目 k1 ガス	k2 降水量	k3 気温	k4 最高気温	k5 最低気温	k6 風速	k7 日照量	k8 旅行日数
2004.9	8.93	8.5	5.1	18.8	24.2	14.9	1.6	4.7	4
2004.10	10.57	9.2	9.3	10.9	16.3	6.5	1.6	4.5	3
2004.11	13.39	12.3	1.8	6.8	13.0	1.8	1.6	6.0	3
2004.12	12.75	6.6	2.2	1.1	7.2	-3.8	1.8	6.3	18
2005.1	17.18	13.2	1.3	-3.7	1.5	-9.0	1.8	5.4	5
2005.2	17.43	14.2	1.1	-2.7	2.1	-7.5	2.0	5.6	0
2005.3	15.30	15	2.8	0.9	6.5	-4.7	2.3	6.1	2
2005.4	11.52	15	1.5	8.8	16.1	1.7	2.2	7.3	0
2005.5	9.50	9	1.8	12.5	19.6	6.3	2.0	6.9	7
2005.6	9.00	10.6	3.9	18.7	24.7	14.0	1.5	4.9	0
2005.7	8.03	8.2	5.4	21.1	26.5	16.7	1.4	4.4	7
2005.8	8.64	8.5	2.0	21.9	27.7	18.1	1.3	5.0	2

斜めに欠測

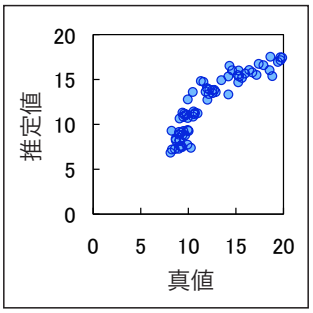
年月	真値 M 電気使用量	測定項目 k1 ガス	k2 降水量	k3 気温	k4 最高気温	k5 最低気温	k6 風速	k7 日照量	k8 旅行日数
2004.9	8.93	8.5	5.1	18.8	24.2	14.9	1.6	4.7	4
2004.10	10.57	9.2	9.3	10.9	16.3	6.5	1.6	4.5	3
2004.11	13.39	12.3	1.8	6.8	13.0	1.8	1.6	6.0	3
2004.12	12.75	6.6	2.2	1.1	7.2	-3.8	1.8	6.3	18
2005.1	17.18	13.2	1.3	-3.7	1.5	-9.0	1.8	5.4	5
2005.2	17.43	14.2	1.1	-2.7	2.1	-7.5	2.0	5.6	0
2005.3	15.30	15	2.8	0.9	6.5	-4.7	2.3	6.1	2
2005.4	11.52	15	1.5	8.8	16.1	1.7	2.2	7.3	0
2005.5	9.50	9	1.8	12.5	19.6	6.3	2.0	6.9	7
2005.6	9.00	10.6	3.9	18.7	24.7	14.0	1.5	4.9	0
2005.7	8.03	8.2	5.4	21.1	26.5	16.7	1.4	4.4	7
2005.8	8.64	8.5	2.0	21.9	27.7	18.1	1.3	5.0	2

市松模様で欠測

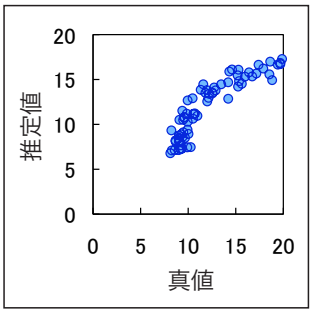
年月	真値 M 電気使用量	測定項目 k1 ガス	k2 降水量	k3 気温	k4 最高気温	k5 最低気温	k6 風速	k7 日照量	k8 旅行日数
2004.9	8.93	8.5	5.1	18.8	24.2	14.9	1.6	4.7	4
2004.10	10.57	9.2	9.3	10.9	16.3	6.5	1.6	4.5	3
2004.11	13.39	12.3	1.8	6.8	13.0	1.8	1.6	6.0	3
2004.12	12.75	6.6	2.2	1.1	7.2	-3.8	1.8	6.3	18
2005.1	17.18	13.2	1.3	-3.7	1.5	-9.0	1.8	5.4	5
2005.2	17.43	14.2	1.1	-2.7	2.1	-7.5	2.0	5.6	0
2005.3	15.30	15	2.8	0.9	6.5	-4.7	2.3	6.1	2
2005.4	11.52	15	1.5	8.8	16.1	1.7	2.2	7.3	0
2005.5	9.50	9	1.8	12.5	19.6	6.3	2.0	6.9	7
2005.6	9.00	10.6	3.9	18.7	24.7	14.0	1.5	4.9	0
2005.7	8.03	8.2	5.4	21.1	26.5	16.7	1.4	4.4	7
2005.8	8.64	8.5	2.0	21.9	27.7	18.1	1.3	5.0	2



総合推定のＳＮ比 = – 3.83 db
相関係数 R = 0.90



総合推定のＳＮ比 = – 3.93 db
相関係数 R = 0.90



総合推定のＳＮ比 = – 3.98 db
相関係数 R = 0.89

図 10　故意にデータを欠測させた場合の比較

$$\eta = \frac{\frac{1}{r}\left(S_\beta - V_e\right)}{V_e} \qquad (1)$$

$$V_e = \frac{1}{l-1}\left(S_T - S_\beta\right) \qquad (2)$$

定項目毎にデータ数が異なる場合、データ数が重み付けの SN 比 η に影響を与えると考えられる。このことが、T 法において欠測データを扱う上での問題点となる。つまり、推定精度（総合推定の SN 比、あるいは相関係数）を悪化させる原因になるのではないかという懸念である。そこで、電気使用量の事例を用いて、その影響度合いを検討してみることにした。

3.2 電気使用量の事例で検証

図 11 で検証の手順を説明する。まず電気使用量の 4 つの測定項目（ガス使用量、降水量、平均気温、平均最高気温）について、それぞれ全く同じデータを用意し、項目数を 2 倍にする。次に、12 ヶ月分のデータについても、それぞれ全く同じデータを用意し、データ数を 2 倍（24 ヶ月分）とする。ここで、測定項目のガス使用量について詳しく説明すると、

年月	真値 M 電気使用量	測定項目 k1 ガス	k1' ガス	k2 降水量	k2' 降水量	k3 気温	k3' 気温	k4 最高気温	k4' 最高気温
2004.9	8.93	8.5	8.5	5.1	5.1	18.8	18.8	24.2	24.2
2004.10	10.57	9.2	9.2	9.3	9.3	10.9	10.9	16.3	16.3
2004.11	13.39	12.3	12.3	1.8	1.8	6.8	6.8	13.0	13.0
2004.12	12.75	6.6	6.6	2.2	2.2	1.1	1.1	7.2	7.2
2005.1	17.18	13.2	13.2	1.3	1.3	-3.7	-3.7	1.5	1.5
2005.2	17.43	14.2	14.2	1.1	1.1	-2.7	-2.7	2.1	2.1
2005.3	15.30	15	15	2.8	2.8	0.9	0.9	6.5	6.5
2005.4	11.52	15	15	1.5	1.5	8.8	8.8	16.1	16.1
2005.5	9.50	9	9	1.8	1.8	12.5	12.5	19.6	19.6
2005.6	9.00	10.6	10.6	3.9	3.9	18.7	18.7	24.7	24.7
2005.7	8.03	8.2	8.2	5.4	5.4	21.1	21.1	26.5	26.5
2005.8	8.64	8.5	8.5	2.0	2.0	21.9	21.9	27.7	27.7
2004.9	8.93	8.5	8.5	5.1	5.1	18.8	18.8	24.2	24.2
2004.10	10.57	9.2	9.2	9.3	9.3	10.9	10.9	16.3	16.3
2004.11	13.39	12.3	12.3	1.8	1.8	6.8	6.8	13.0	13.0
2004.12	12.75	6.6	6.6	2.2	2.2	1.1	1.1	7.2	7.2
2005.1	17.18	13.2	13.2	1.3	1.3	-3.7	-3.7	1.5	1.5
2005.2	17.43	14.2	14.2	1.1	1.1	-2.7	-2.7	2.1	2.1
2005.3	15.30	15	15	2.8	2.8	0.9	0.9	6.5	6.5
2005.4	11.52	15	15	1.5	1.5	8.8	8.8	16.1	16.1
2005.5	9.50	9	9	1.8	1.8	12.5	12.5	19.6	19.6
2005.6	9.00	10.6	10.6	3.9	3.9	18.7	18.7	24.7	24.7
2005.7	8.03	8.2	8.2	5.4	5.4	21.1	21.1	26.5	26.5
2005.8	8.64	8.5	8.5	2.0	2.0	21.9	21.9	27.7	27.7
データ数		12	24	12	24	12	24	12	24
重み付けの SN 比 η		0.8	0.8	0.3	0.3	10	10	13	13

データ数が異なっても、同じ SN 比の値になって欲しい

図 11　電気使用量の事例で検証

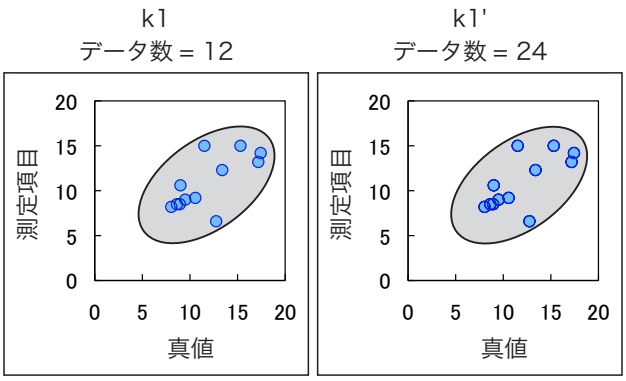


図 12　k1 と k1' をグラフで説明

k_1 および k_1' は、共にデータ数 =24 の同じデータであるが、k_1 の下半分（灰色：12 ヶ月分）のデータを欠測値として設定することにする。

降水量 k_2 および k_2'、平均気温 k_3 および k_3'、平均最高気温 k_4 および k_4' についても、同様の設定をする。

k_1 および k_1' は、全く同じデータをダブらせたものであるから、そのバラツキの大きさ、つまり重み付けの SN 比 η は、同じ値になるのが理想である。

説明がわかりにくいかもしれないので、図 12 で同じことを説明する。

左の k_1 に対して、右の k_1' は k_1 のデータを 2 倍にしてダブらせたものである。k_1' のグラフデータは、重なって見えるが、データ数は k_1 の 2 倍ある。両者のバラツキの大きさは全く同じなので、SN 比の値は同じになるのが理想である。

図 13 に、従来のゼロ点比例式で算出した k_x および k_x' の重み付けの SN 比 η を示す。全く同じバラツキにもかかわらず、SN 比 η は異なる値となっている。これはデータ数が SN 比 η に影響を与えているからである。この差が、推定精度（総合推定の SN 比および相関係数）に悪影響を与える可能性がある。

4. 欠測データ数に依存しない SN 比の検討

4.1 エネルギー比型 SN 比で検証

欠測データ数が重み付けの SN 比 η に与える影響を無くすため、清水らが提案するエネルギー比型 SN 比 [2] を採用してみることにした。

$$\eta_E = \frac{S_\beta}{S_e} \qquad (3)$$

式 (3) に示すように、エネルギー比型 SN 比 η_E は、データ数 l に依存しない。よって、今回の欠測データにおける T 法で採用するには好都合といえる。

図 14 に、エネルギー比型 SN 比での検証結果を示す。エネルギー比型 SN 比で算出した k_x および k_x' の重み付けの SN 比 η_E は、全く同じ値となった。

エネルギー比型 SN 比を T 法に採用した場合、どれくらい推定精度が向上するのかを確かめるため、2 つの事例で検証することにする。

4.2 適用事例その 1「我が家の電気使用量の解析」

図 15 に、左上のデータを故意に欠測せた場合の推定結果を示す。従来の SN 比とエネルギー比型 SN 比を比較すると、ほぼ同じ推定精度となった。また、図 16 に示すように、右下のデータを故意に欠測せた場合も、ほぼ同じ推定精度となった。

4.3 適用事例その 2「世界各国の出生率の解析」

アメリカ中央情報局（CIA）の Web サイトからデータを引用し、世界各国の出生率の解析を行った。既知データとして、15 カ国（Afghanistan、Albania、Algeria など）のデータを用い、未知データとして 133 カ国の出生率を推定した。

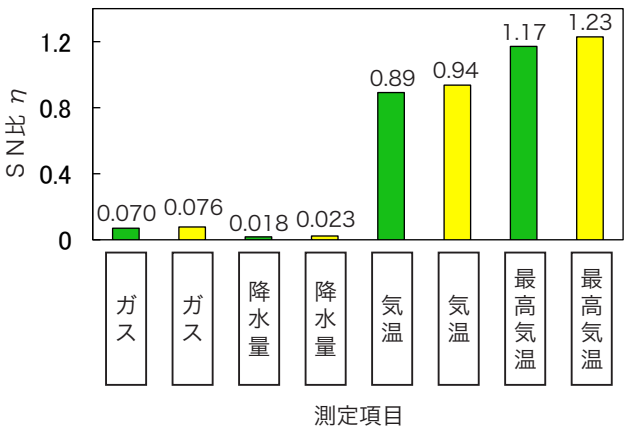


図 13　従来のゼロ点比例式での検証結果

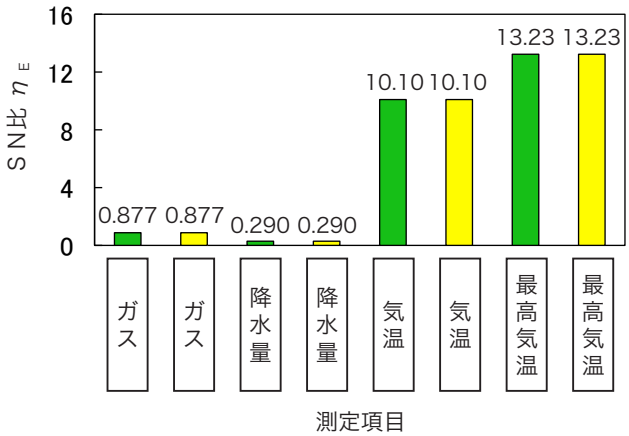
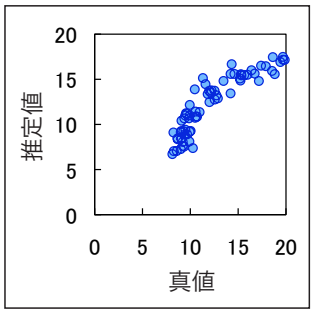


図 14　エネルギー比型 SN 比での検証結果

従来の SN 比

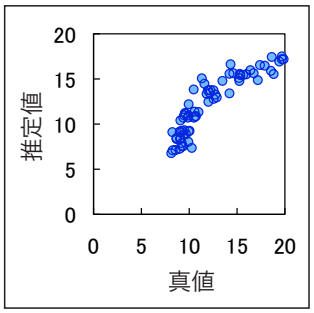


総合推定のＳＮ比＝－3.74 db
相関係数 R = 0.90

左上を欠測

	真値	測定項目							
	M	k1	k2	k3	k4	k5	k6	k7	k8
年月	電気使用量	ガス	降水量	気温	最高気温	最低気温	風速	日照量	旅行日数
2004.9	8.93	8.5	5.1	18.8	24.2	14.9	1.6	4.7	4
2004.10	10.57	9.2	9.3	10.9	16.3	6.5	1.6	4.5	3
2004.11	13.39	12.3	1.8	6.8	13.0	1.8	1.6	6.0	3
2004.12	12.75	6.6	2.2	1.1	7.2	-3.8	1.8	6.3	18
2005.1	17.18	13.2	1.3	-3.7	1.5	-9.0	1.8	5.4	5
2005.2	17.43	14.2	1.1	-2.7	2.1	-7.5	2.0	5.6	0
2005.3	15.30	15	2.8	0.9	6.5	-4.7	2.3	6.1	2
2005.4	11.52	15	1.5	8.8	16.1	1.7	2.2	7.3	0
2005.5	9.50	9	1.8	12.5	19.6	6.3	2.0	6.9	7
2005.6	9.00	10.6	3.9	18.7	24.7	14.0	1.5	4.9	0
2005.7	8.03	8.2	5.4	21.1	26.5	16.7	1.4	4.4	7
2005.8	8.64	8.5	2.0	21.9	27.7	18.1	1.3	5.0	2

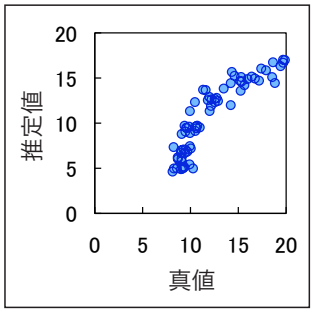
エネルギー比型 SN 比



総合推定のＳＮ比＝－3.71 db
相関係数 R = 0.90

図 15　電気使用量（故意に左上を欠測）

従来の SN 比

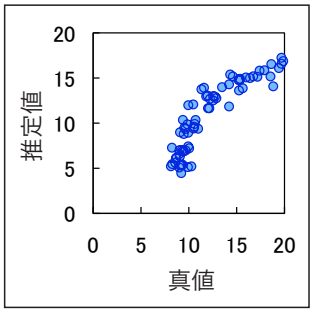


総合推定のＳＮ比＝－6.08 db
相関係数 R = 0.90

右下を欠測

	真値	測定項目							
	M	k1	k2	k3	k4	k5	k6	k7	k8
年月	電気使用量	ガス	降水量	気温	最高気温	最低気温	風速	日照量	旅行日数
2004.9	8.93	8.5	5.1	18.8	24.2	14.9	1.6	4.7	4
2004.10	10.57	9.2	9.3	10.9	16.3	6.5	1.6	4.5	3
2004.11	13.39	12.3	1.8	6.8	13.0	1.8	1.6	6.0	3
2004.12	12.75	6.6	2.2	1.1	7.2	-3.8	1.8	6.3	18
2005.1	17.18	13.2	1.3	-3.7	1.5	-9.0	1.8	5.4	5
2005.2	17.43	14.2	1.1	-2.7	2.1	-7.5	2.0	5.6	0
2005.3	15.30	15	2.8	0.9	6.5	-4.7	2.3	6.1	2
2005.4	11.52	15	1.5	8.8	16.1	1.7	2.2	7.3	0
2005.5	9.50	9	1.8	12.5	19.6	6.3	2.0	6.9	7
2005.6	9.00	10.6	3.9	18.7	24.7	14.0	1.5	4.9	0
2005.7	8.03	8.2	5.4	21.1	26.5	16.7	1.4	4.4	7
2005.8	8.64	8.5	2.0	21.9	27.7	18.1	1.3	5.0	2

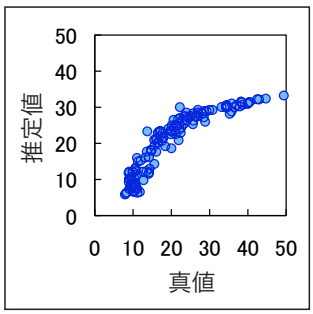
エネルギー比型 SN 比



総合推定のＳＮ比＝－6.17 db
相関係数 R = 0.89

図 16　電気使用量（故意に右下を欠測）

従来の SN 比

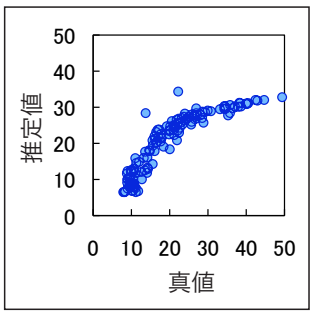


総合推定のＳＮ比＝－13.01 db
相関係数 R = 0.90

左上を欠測

	真値	測定項目									
	M	k1	k2	k3	k4	k5	k6	k7	k8		
国	出生率	人口	年齢の中央値	死亡率	乳児死亡率	GDP/人	失業率	インフレ率	携帯電話普及率	ネット普及率	国防費 GDP 比
Afghanistan	45.82	32738376	17.6	19.56	154.67	1000	0.4	0.13	0.08	0.02	0.019
Albania	15.22	3619778	29.5	5.44	19.31	6300	0.13	0.029	0.42	0.13	0.0149
Algeria	17.03	33769669	26	4.62	28.75	6500	0.13	0.037	0.62	0.07	0.033
Argentina	16.32	40677348	30.3	7.54	13.87	13300	0.141	0.088	0.77	0.20	0.013
Armenia	12.53	2968586	31.1	8.34	20.94	4900	0.071	0.044	0.11	0.06	0.065
Australia	11.90	20600856	37.4	7.62	4.51	36300	0.044	0.023	0.96	0.74	0.024
Austria	8.66	8205533	41.7	9.91	4.48	38400	0.044	0.022	1.13	0.51	0.009
Azerbaijan	17.52	8177717	27.9	8.32	56.43	7700	0.043	0.166	0.41	0.10	0.026
Bahamas	17.06	307451	28.4	9.22	23.67	25000	0.076	0.024	0.74	0.34	0.005
Bahrain	17.26	718306	29.9	4.29	15.64	32100	0.15	0.034	1.25	0.22	0.045
Bangladesh	28.86	153546901	22.8	8	57.45	1300	0.01	0.084	0.12	0.00	0.015
Barbados	12.48	281968	35.4	8.58	11.05	19300	0.107	0.055	0.73	0.57	0.005
Belarus	9.62	9685768	38.4	13.92	6.53	10900	0.016	0.084	0.62	0.57	0.014
Belgium	10.22	10403951	41.4	10.38	4.5	35300	0.075	0.018	0.93	0.46	0.013
Belize	27.84	301270	20.1	5.77	23.65	7900	0.094	0.03	0.39	0.11	0.014

エネルギー比型 SN 比



総合推定のＳＮ比＝－13.48 db
相関係数 R = 0.89

図 17　出生率（故意に左上を欠測）

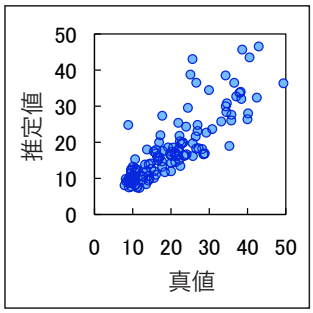
図 17 に示すように、左上のデータを故意に欠測せた場合の推定結果を示す。従来の SN 比とエネルギー比型 SN 比を比較すると、ほぼ同じ推定精度となった。また、図 18 に示すように、右下のデータを故意に欠測せた場合も、ほぼ同じ推定精度となった。

5. さいごに

今回の研究では、T 法における欠測データの活用について、重み付けの SN 比の検討を行った。その結果、従来の SN 比とエネルギー比型 SN 比の差は小さく、推定精度に与える影響は軽微であった。

今回の研究で示したように、T 法は欠測データが存在する場合でも、問題なく解析できる便利な手法

従来の SN 比

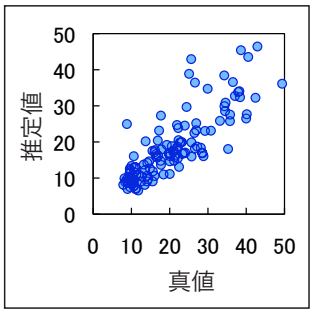


総合推定のＳＮ比 = – 15.21 db
相関係数 R = 0.85

右下を欠測

	真値	測定項目									
	M	k1	k2	k3	k4	k5	k6	k7	k8		
国	出生率	人口	年齢の中央値	死亡率	乳児死亡率	GDP/人	失業率	インフレ率	携帯電話普及率	ネット普及率	国防費GDP 比
Afghanistan	45.82	32738376	17.6	19.56	154.67	1000	0.4	0.13	0.08	0.02	0.019
Albania	15.22	3619778	29.5	5.44	19.31	6300	0.13	0.029	0.42	0.13	0.0149
Algeria	17.03	33769669	26	4.62	28.75	6500	0.13	0.037	0.62	0.07	0.033
Argentina	16.32	40677348	30.3	7.54	13.87	13300	0.141	0.088	0.77	0.20	0.013
Armenia	12.53	2968586	31.1	8.34	20.94	4900	0.071	0.044	0.11	0.06	0.065
Australia	11.90	20600856	37.4	7.62	4.51	36300	0.044	0.023	0.96	0.74	0.024
Austria	8.66	8205533	41.7	9.91	4.48	38400	0.044	0.022	1.13	0.51	0.009
Azerbaijan	17.52	8177717	27.9	8.32	56.43	7700	0.043	0.166	0.41	0.10	0.026
Bahamas	17.06	307451	28.4	9.22	23.67	25000	0.076	0.024	0.74	0.34	0.005
Bahrain	17.26	718306	29.9	4.29	15.64	32100	0.15	0.034	1.25	0.22	0.045
Bangladesh	28.86	153546901	22.8	8	57.45	1300	0.01	0.084	0.12	0.00	0.015
Barbados	12.48	281968	35.4	8.58	11.05	19300	0.107	0.055	0.73	0.57	0.005
Belarus	9.62	9685768	38.4	13.92	6.53	10900	0.016	0.084	0.62	0.57	0.014
Belgium	10.22	10403951	41.4	10.38	4.5	35300	0.075	0.018	0.93	0.46	0.013
Belize	27.84	301270	20.1	5.77	23.65	7900	0.094	0.03	0.39	0.11	0.014

エネルギー比型 SN 比



総合推定のＳＮ比 = – 15.39 db
相関係数 R = 0.84

図 18　出生率（故意に右下を欠測）

であり、従来の重回帰分析にはないメリットを有している。今後、技術者を支援するツールのひとつとして、T 法を普及していきたいと考えている。

謝　辞

T 法における欠測データの活用方法をご教示いただいた (財) 長野県テクノ財団の岩下幸廣氏に、感謝の意を表する。

■参考文献

1) 田口玄一：目的機能と基本機能 (6), 品質工学 ,Vol.13,No.3,pp.309-314,2005
2) 清水豊他：新 SN 比の研究 (4), 第 16 回品質工学会研究発表大会論文集 ,(2008),pp.422-425