

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2019年10月3日(03.10.2019)



(10) 国際公開番号

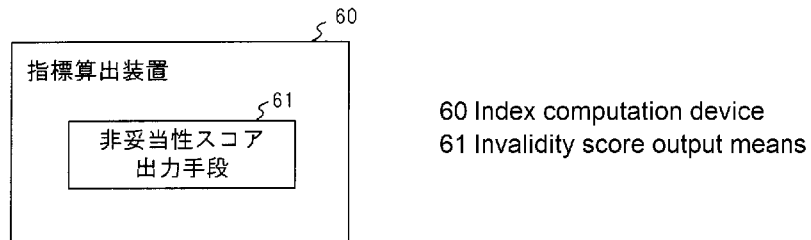
WO 2019/187341 A1

- (51) 国際特許分類:
G06Q 10/04 (2012.01)
- (21) 国際出願番号: PCT/JP2018/043910
- (22) 国際出願日: 2018年11月29日(29.11.2018)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2018-067622 2018年3月30日(30.03.2018) JP
- (71) 出願人: NECソリューションイノベータ株式会社(NEC SOLUTION INNOVATORS, LTD.) [JP/JP]; 〒1368627 東京都江東区新木場一丁目18番7号 Tokyo (JP).
- (72) 発明者: 玉野 浩嗣 (TAMANO Hiroshi); 〒1368627 東京都江東区新木場一丁目18番7号 NECソリューションイノベータ株式会社内 Tokyo (JP). 谷口 敦(TANIGUCHI Atsushi); 〒1368627 東京都江東区新木場一丁目18番7号 NECソリューションイノベータ株式会社内 Tokyo (JP). 清家 勝代(SEIKE Katsuyo); 〒1368627 東京都江東区新木場一丁目18番7号 NECソリューションイノベータ株式会社内 Tokyo (JP). 田中 博典(TANAKA Hirofumi); 〒1368627 東京都江東区新木場一丁目18番7号 NECソリューションイノベータ株式会社内 Tokyo (JP). 田尻 俊宗(TAJIRI Toshikazu); 〒1368627 東京都江東区新木場一丁目18番7号 NECソリューションイノベータ株式会社内 Tokyo (JP).

(54) Title: INDEX COMPUTATION DEVICE, PREDICTION SYSTEM, PROGRESS PREDICTION EVALUATION METHOD, AND PROGRAM

(54) 発明の名称: 指標算出装置、予測システム、経過予測評価方法およびプログラム

【図11】



(57) Abstract: An objective of the present invention is to automatically determine the invalidity of sequential data including a prediction value obtained on the basis of a hypothetical condition. Provided is an index computation device 60 comprising an invalidity score output means 61 for, when sequential data of a prescribed item to be predicted is inputted thereto, outputting an invalidity score which is an index indicating the invalidity of the sequential data. The sequential data includes three or more instances of data indicating a value of the item to be predicted in association with time, and at least one of the three or more instances of data indicates a prediction value. The invalidity score is based on the difference between the sequential data and a curvilinear model obtained by fitting the sequential data to a prescribed function form.

(57) 要約: 仮定の条件に基づいて得られる予測値を含む系列データの非妥当性を自動で判断する。指標算出装置60は、所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データが入力されると、系列データの非妥当性を示す指標である非妥当性スコアであって、系列データを所定の関数形にフィッティングして得られる曲線モデルと系列データとの誤差に基づく非妥当性スコアを出力する非妥当性スコア出力手段61を備える。

(74) 代理人:岩壁 冬樹, 外(IWAKABE Fuyuki et al.);
〒1040031 東京都中央区京橋二丁目 8 番
7 号 読売八重洲ビル 6 階 サンライズ
国際特許事務所 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保
護が可能): AE, AG, AL, AM, AO, AT, AU, AZ,
BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH,
CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保
護が可能): ARIPO (BW, GH, GM, KE, LR, LS,
MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM,
ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ,
TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ,
DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT,
LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS,
SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 一 国際調査報告 (条約第21条(3))

明 細 書

発明の名称：

指標算出装置、予測システム、経過予測評価方法およびプログラム

技術分野

[0001] 本発明は、時間の経過にともなって値が変化する推移グラフ等における系列データの妥当性を判断するための指標算出装置、それを用いた予測システムに関する。また、本発明は、時間の経過にともなって値が変化する変数の2以上の時点における予測値を含む経過予測の妥当性を評価する経過予測評価方法および経過予測評価プログラムに関する。

背景技術

[0002] 例えば、健康診断等で測られる従業員等の検査値やそれに基づく生活習慣病の発症確率等の経年変化を予測して、従業員の各々に健康に関するアドバイスをすることを考える。具体的には、過去の健康診断結果やそのときの生活習慣を示すデータを基に、今の生活習慣を3年間続けた場合の将来の状態（検査値の経年変化や発症確率等）を予測して、産業医、保険師等が生活習慣の見直し等を提案（保健指導）したり、従業員自身がセルフチェックしたりすることを考える。

[0003] その場合、予測値の推移を得るための方法として次の方法が考えられる。まず、過去のデータから1年先の予測値を得る予測モデルを学習する。例えば、過去の予測対象の実績値（検査値）と対応づけて、それと相関し得るさらに過去の検査値や予測対象者の属性（年齢等）やそのときの生活習慣などを示す訓練データを用いて、1年後の予想対象項目を目的変数とし、それに相関し得る他の項目を説明変数とする予測モデルを学習する。そして、得られた予測モデルに対して、予測対象の値を得る時点（予測時点）を変えながら説明変数を入力して1年先の予測値を得る処理を、数年分繰り返す。このとき、説明変数のうち生活習慣に関する項目を一定とすることで、今の生活習慣を3年間続けた場合の予測値の推移を得ることができる。

- [0004] このような予測値の推移に基づいて保健指導やセルフチェックを行えば、より効果的かつ効率的な、指導対象者や自身の不調予防・健康増進およびそのための行動変容が見込まれる。
- [0005] 健康予測や健康支援に関する技術として、非特許文献1には、健康診断の成績とライフスタイルを元に健康危険度を予知する健康危険度予測システムの例が記載されている。非特許文献1に記載の健康危険度予測システムは、検査値予測システムと、発症予測システムの2つのサブシステムを含む。検査値予測システムは、現時点で軽度の検査異常を伴っていてかつライフスタイルに問題がある人を対象に、ライフスタイルの改善に伴う検査成績の改善度合いを示すものである。また、発症予想システムは、現時点で検査異常を伴わないがライフスタイルに偏りがある人を対象に、望ましくないライフスタイルを継続した場合とライフスタイルを改善した場合の疾病発症率を予測するものである。非特許文献1によれば、いずれのサブシステムによる健康支援も医療従事者と受診者との間でインタラクティブに行われ、受診者は自分の行動変容の効果を認識できるとされている。
- [0006] しかし、単純に予測モデルを用いてこのような仮定における将来の予測値の推移を得る場合、得られる予測値の推移が一般的な知見とは異なる変化を示す場合があった。
- [0007] 予測値の検証に関して、例えば、特許文献1には、需要予測モデルの評価方法が記載されている。特許文献1に記載の方法は、需要予測モデルの評価期間における需要実績値と需要予測値を取り込み、製品の供給サイクルごとに需要実績値と需要予測値との乖離値を算出し、算出された乖離値に基づいて需要予測モデルを評価する。

先行技術文献

特許文献

- [0008] 特許文献1：特開平10-228463号公報

非特許文献

- [0009] 非特許文献1：高橋 英孝、吉田 勝美、「ライフスタイル改善のためのHR

A（健康危険度予測システム）」、日本人間ドック学会誌（JHD）Vol. 11 No. 4 1997年、p. 123-128.

発明の概要

発明が解決しようとする課題

[0010] しかし、非特許文献1に記載の予測システムでは、予測値の推移の妥当性について、何ら考慮されていない。

[0011] また、特許文献1に記載の方法は、実績値を用いて過去の予測値の総合評価を行うものであり、少なくとも予測をおこなった時点での実績値が得られることを前提としている。しかし、上記のような仮定を用いて将来の複数時点の値を予測する経過予測では、その時点において仮定した条件が全て合致するとは限らない。すなわち、予測時の条件と全て合致した実績値が必ずしも得られない。また、予測結果を表示する前にその妥当性を自動で判断する用途等には、実績値を得た後に判断を行う特許文献1の方法は適用できない。

[0012] なお、似た属性をもつ他の人の過去の実績値から同じ条件のものを探して実績値の代わりにすることも考えられる。しかし、全ての対象者について、過去の検査値と仮定した条件とを含む多くの説明変数の値が全ての予測時点で合致するものが見つかるとは限らない。そもそも個体差が生じやすい検査値を予測する上記の経過予測において、ごく少数の個別的な実績値を基に仮の条件での将来の予測値の推移の妥当性を判断しても、精度のよい判断結果を得ることは困難である。

[0013] 課題は、仮定の条件に基づいて得られる予測値の推移の妥当性を判断できる客観的な指標がなく、予測結果が妥当であるか否かや、妥当でない場合にどれくらい妥当でないかを自動で判断するのが困難な点である。なお、ここで「妥当でない」とは、少なくとも予測値を扱う人（上記の例でいえば、保健指導を行う産業医、保険師等や、セルフチェックを行う本人等）の知見において不確か（説明不能）なことをいう。

[0014] 例えば、生活習慣に関する説明変数を一定にしたときの生活習慣病の発症

確率やそれに関連する検査値を予測して得られた予測値の推移を基にアドバイスをを行うことを考える。

[0015] 一般的な感覚によれば、生活習慣を一定にすると、例えば図13(a)に示すように、生活習慣病の発症確率やそれに関連する検査値の推移は何かの値に漸近していくのが自然である。

[0016] しかし、単に、所定の予測時間単位における1つ先の時点（例えば、1年後）の予測値を予測する予測モデルを繰り返し適用して経過予測をした場合、生活習慣を一定にしているにも関わらず、図13(b)に示すように、ある年では上向きに変化し、別のある年では下向きに変化するなど、変化の向き（プラスかマイナスか）が定まらずに乱高下するようなガタついたグラフ形状となる予測結果が示される可能性がある。このような予測結果を基に、生活習慣の改善のアドバイスをしようとしても、アドバイスをする側・受ける側の双方において違和感が生じてしまうため、アドバイスの根拠として使用できないという問題がある。

[0017] 例えば上記の健康シミュレーションのように、過去の健康診断結果やそのときの生活習慣を示すデータを基に経過予測を行った結果を基に、産業医、保険師等が生活習慣の見直し等を提案（保健指導）したり、従業員自身がセルフチェックしたりすることを考えた場合、予測機構としては、いかに上記のような解釈できないような予測結果が出力されることを防止するかが重要となる。

[0018] そのためには、仮定の条件に基づいて得られる予測値を含む系列データを得た際に、実績値を待たずに、得られた系列データで示される予測値の推移が妥当であるか否か、または妥当でない場合にはどれくらい妥当でないかを表す客観的な指標が得られることが望ましい。

[0019] 例えば、予測結果の出力判断をドメインエキスパートによる目視で行う方法では、コストがかかるだけでなく、出力までに時間がかかるという問題がある。また、いくつかある予測モデルの候補の中から上記のような解釈できないモデルを選択する場合においても、ドメインエキスパートによる目視で

行う方法では同様の問題がある。

[0020] 本発明は、上述した課題に鑑みて、仮定の条件に基づいて得られる予測値を含む系列データの非妥当性を自動で判断できる指標算出装置、予測システム、経過予測評価方法および経過予測評価プログラムを提供することを目的とする。

課題を解決するための手段

[0021] 本発明による指標算出装置は、所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データが入力されると、系列データの非妥当性を示す指標である非妥当性スコアであって、系列データを所定の関数形にフィッティングして得られる曲線モデルと系列データとの誤差に基づく非妥当性スコアを出力する非妥当性スコア出力手段を備えることを特徴とする。

[0022] また、本発明による予測システムは、所定の予測対象項目について、学習済みの予測モデルを用いて所定の予測時点における予測値を得るとともに、得られた予測値を含む系列データであって時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データを生成する予測手段と、系列データの非妥当性を示す指標である非妥当性スコアであって、系列データを所定の関数形にフィッティングして得られる曲線モデルと系列データとの誤差に基づく非妥当性スコアを算出する非妥当性スコア算出手段と、非妥当性スコアに基づいて、系列データ、系列データに含まれる予測値または予測値を得た予測モデルを評価する評価手段とを備えることを特徴とする。

[0023] また、本発明による経過予測評価方法は、情報処理装置が、所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データが入力されると、系列データの非妥当性を示す指標である非妥当性スコアであって、系列データを所定の関数形にフィッティングして得

られる曲線モデルと系列データとの誤差に基づく非妥当性スコアを算出し、非妥当性スコアに基づいて、系列データまたは系列データに含まれる予測値もしくは予測値を得た予測モデルを評価することを特徴とする。

[0024] また、本発明による経過予測評価プログラムは、コンピュータに、所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データが入力されると、系列データの非妥当性を示す指標である非妥当性スコアであって、系列データを所定の関数形にフィッティングして得られる曲線モデルと系列データとの誤差に基づく非妥当性スコアを算出する処理、および非妥当性スコアに基づいて、系列データまたは系列データに含まれる予測値もしくは予測値を得た予測モデルを評価する処理を実行させることを特徴とする。

発明の効果

[0025] 本発明によれば、仮定の条件に基づいて得られる予測値を含む系列データの非妥当性を自動で判断できる。

図面の簡単な説明

[0026] [図1]第1の実施形態の指標算出装置の構成例を示すブロック図である。
[図2]非妥当性スコアの概要を示す説明図である。
[図3]系列データのグラフ表示の例を示す説明図である。
[図4]x__scaleの算出方法の概略を示す説明図である。
[図5]第1の実施形態の指標算出装置の動作の一例を示すフローチャートである。
[図6]第2の実施形態の予測システムの構成例を示すブロック図である。
[図7]第2の実施形態の予測システムの動作の一例を示すフローチャートである。
[図8]第3の実施形態の予測システムの構成例を示すブロック図である。
[図9]第3の実施形態の予測システムの動作例を示すフローチャートである。
[図10]本発明の実施形態にかかるコンピュータの構成例を示す概略ブロック

図である。

[図11]本発明の指標算出装置の概要を示すブロック図である。

[図12]本発明の予測システムの概要を示すブロック図である。

[図13]予測値の推移例を示すグラフである。

発明を実施するための形態

[0027] 実施形態 1.

以下、図面を参照して本発明の実施形態について説明する。図 1 は、本実施形態の指標算出装置の構成例を示すブロック図である。図 1 に示す指標算出装置 10 は、スコア算出部 11 を備える。

[0028] スコア算出部 11 は、時間と対応づけて予測対象項目の値を示す 3 つ以上のデータであってそのうちの少なくとも 1 つが予測値を示すデータを含む系列データを入力とし、その系列データの非妥当性を表す指標である非妥当性スコアを算出して出力する。

[0029] 以下では、非妥当性スコアを、入力された系列データが所定の漸近モデルとどれだけ離れているかを示す指標として算出する。

[0030] 既に説明したように、ある生活習慣を継続したときの検査値の予測値は何かの値に漸近するのが自然である。そこで、スコア算出部 11 は、入力された系列データと所定の漸近モデルとの誤差に基づいて、入力された予測値系列の説明不能度合い（非妥当性スコア）を算出する（図 2 参照）。なお、図 2 は、非妥当性スコアの概要を示す説明図であり、図 2（a）は非妥当性スコアが小さい場合の例、図 2（b）は非妥当性スコアが大きい場合の例である。なお、図 2 において網かけの領域はその検査値の正常範囲を表している。

[0031] ここで、漸近モデルは、時間を X 軸、予測項目を Y 軸としたとき、X 軸と平行な漸近線が存在する曲線を表す曲線モデル、より具体的には、 $x \rightarrow \infty$ のとき $y(x)$ がある値に収束する関数で表される曲線モデルをいう。ここで、 x は、系列データにおいて各データが対応する時間軸上の点（座標）を表し、 $y(x)$ は x 時点における予測項目値を表す。なお、漸近モデルは、以

下の式（１）で表される条件を少なくとも満たす関数で表される曲線モデルであってもよい。ここで、 a は任意の定数である。なお、存在する漸近線は１つに限らず、例えば、ロジスティック関数や逆正接関数と呼ばれる関数のように２つの漸近線が存在する関数で表されるものも含む。

[0032] [数1]

$$\lim_{x \rightarrow \infty} y(x) = a \quad \dots \quad (1)$$

[0033] スコア算出部 11 は、漸近モデルとして、予め定めた１つの関数形で表される曲線モデルを用いることもできるが、例えば、入力された系列データを、上記の条件を満たす所定の２以上の関数形にフィッティングして得られるモデルを用いてもよい。

[0034] フィッティングは、例えば、以下の式（２）に示すような予め定められた損失関数を最小化するモデルパラメータ θ の解（ハット付き θ ）を探索することによって行ってもよい。

[0035] [数2]

$$\hat{\theta} = \operatorname{argmin}_{\theta} \sum_n \operatorname{loss}(f(x_n, \theta), y_n) \quad \dots \quad (2)$$

$ex)$

$$\operatorname{loss}(y_1, y_2) = (y_1 - y_2)^2,$$

$$f(x, \theta) = c + ba^x$$

$$\text{s.t. } \theta = \{a, b, c\}, 0 < a < 1$$

[0036] 式（２）において、 n はフィッティングを行う値の時点を表し、 $\operatorname{loss}()$ は誤差関数を表し、 $f()$ はフィッティング先の関数形を表す。なお、 $f(x_n, \theta)$ で、その関数形 $f()$ に任意の時点 x_n と１組のモデルパラメータ θ とを与えたときの出力を表し、 $f(x_n, \hat{\theta})$ で、フィッティングにより得られた漸近モデルにおける任意の時点 x_n の出力を表す。式（２）に示す例では、誤差関数として二乗損失を用いているが、誤差関数は二乗損失に

限らない。

[0037] スコア算出部 11 は、例えば、このようにして得られた漸近モデルと、入力された予測値の系列データとの誤差を算出し、非妥当性スコアとして出力してもよい。

[0038] スコア算出部 11 は、例えば、以下の式 (3) で示される誤差値 (error) を非妥当性スコアとして出力してもよい。

[0039] [数3]

$$error = \sum_n loss(f(x_n, \hat{\theta}), y_n) \quad \dots \quad (3)$$

[0040] スコア算出部 11 は、フィッティングに用いるデータと、誤差の算出に用いるデータとを独立に指定してもよい。なお、スコア算出部 11 は、それらの指定をユーザから受け付けることも可能である。このとき、フィッティングに用いるデータ (第 1 グループに属するデータ) と、誤差の算出に用いるデータ (第 2 グループに属するデータ) とは完全に一致していなくてもよい。

[0041] 一例として、N 個のデータを含む系列データが入力された場合に、フィッティングを前半の N' (ただし N' < N) 個のデータを用いて行い、誤差の計算を残りのデータ (N - N' 個) や全てのデータ (N 個) の値を用いて行ってもよい。また、他の例として、フィッティングを 1 番目、3 番目、5 番目のデータなど系列データにおいて連続していない時点のデータを用いて行い、誤差の計算を全てのデータを用いて行う、なども可能である。

[0042] 例えば、5 個のデータを含む系列データが入力された場合に、スコア算出部 11 は、例えば次のようにしてフィッティングおよび誤差の計算を行ってもよい。

[0043] [フィッティングと誤差計算のパターン例]

・前半 3 個のデータでフィッティングを行い、後半 2 個のデータで誤差を計算する。

- ・前半3個のデータでフィッティングを行い、全てのデータで誤差を計算する。

- ・前から1、3、5番目のデータでフィッティングを行い、全てのデータで誤差を計算する。

[0044] ここで、スコア算出部11に入力される系列データに含まれるデータ数（予測値の数） N_p は特に限定されないが、少なくとも1つは含まれるものとする。なお、実際は、系列データに、経過予測を行った時点数分の予測値を示すデータが少なくとも含まれることが好ましい。なお、系列データは、過去の実績値を示すデータを含んでいてもよく、その場合、上記の N は、過去の実績値を示すデータを含めた総データ数を表すものとする。なお、 N は3以上を前提とするが、フィッティング精度の観点からは、例えば4以上がより好ましい。

[0045] なお、系列データに実績値を示すデータが含まれている場合は、誤差の計算を、予測値を示すデータのみを用いて行ってもよい。

[0046] また、スコア算出部11は、フィッティングを行う前に、 X 軸の値である x 座標（各予測項目の値が対応する時間を表す値）をリスケールしてもよい。図3（a）に示すように、系列データをグラフ表示する際、縦軸（ Y 軸：予測値）の目盛の単位が50、横軸（ X 軸：時間）の目盛の単位が1など、縦軸と横軸とで数値の単位（目盛の幅）が大きく異なることが考えられる。その場合、そのままの数値を用いてフィッティングしても、見た人が期待するような漸近モデルが得られない。これは、図3（b）に示すように、 X 軸と Y 軸とで尺度が大きく異なる値を等間隔でグラフ表示したときのグラフ形状でフィッティングが行われるために、実際に表示させたときに妥当なグラフ形状とされる曲線とは異なる形状の曲線にフィッティングされてしまうためである。

[0047] このような不都合を解消するために、図3（b）に示すように、実際の表示に合わせて、系列データに含まれる各データに対応づけられている時間の値（ x 座標）をリスケールした上でフィッティングを行うのが好ましい。

[0048] 例えば、スコア算出部 11 は、入力された系列データを表示するグラフの目盛設定等の表示に関するパラメータ（表示パラメータ）が得られる場合には、X軸の主目盛の幅がY軸の主目盛の幅と同じ単位（図3（c）の例でいうと、50単位）になるように、x座標を変換してもよい。この場合、X軸の主目盛の幅の単位=1に対して、Y軸の主目盛の幅の単位=50（50倍）となっているため、x座標も50倍とする。以下、このようなx軸のリスケール時の倍率を、リスケールパラメータ x_scale という場合がある。

[0049] また、スコア算出部 11 は、リスケールパラメータ x_scale を次のようにして求めることも可能である。スコア算出部 11 は、系列データとともに、その系列データを表示する際のパラメータである表示パラメータを入力し、系列データから得られる情報と、該表示パラメータとに基づいて、リスケールパラメータ x_scale を計算することも可能である。

[0050] 以下の式（4）は、リスケールパラメータ x_scale の算出式の一例である。式（4）において、 y_{max} 、 y_{min} は系列データに含まれる予測項目の最大値、最小値を表す。また、 N_d は系列データに含まれるデータ数（表示させる予測項目の点数）を表す。また、 A_r はその系列データの表示グラフの縦横比 A_r （すなわち縦幅に対する横幅の比率）を表す。なお、式（4）において、0.8は縦方向の表示比率、0.9は横方向の表示比率を表しているが、これらの値は適宜調整される。

$$[0051] \quad x_scale = ((y_{max} - y_{min}) / 0.8 * A_r * 0.9) / (N_d - 1) \quad \dots (4)$$

[0052] 式（4）に示す例では、 y_{max} 、 y_{min} および N_d が系列データから得られる情報に相当し、 A_r 、縦方向の表示比率および横方向の表示比率が表示パラメータに相当する。

[0053] また、図4は、式（4）で示した x_scale の算出方法の概略を示す説明図である。例えば、縦横比が1:2（ $A_r = 2$ ）、枠内に5点を表示す

る場合、 x_scale は次のように算出される。

$$[0054] \quad x_scale = ((y_{max} - y_{min}) / 0.8 * 2 * 0.9) / 4$$

[0055] 上記の例では、系列データに含まれる各データに対応づけられている時間の単位が予測単位（すなわち予測時点が1つ増えるごとに1増える数）であることを前提として、 x_scale を、 y 軸の単位で表現した各データの x 方向の間隔を表す指標として算出している。したがって、系列データにおける時間の単位が予測単位時間以外の場合や、表示する際の x 方向の間隔の単位が1以外の場合には、各データに対応づけられている x 座標を予測単位時間や x 方向の間隔の単位で割って x 軸の単位を1とした上で、 x_scale 倍すればよい。

[0056] なお、スコア算出部11は、 x_scale の指定を受けつけることも可能である。例えば、スコア算出部11は、系列データとともに、 x_scale を入力してもよい。なお、スコア算出部11は、系列データとともに、上述したような表示パラメータを入力して x_scale を算出することも可能である。

[0057] 次に、本実施形態の動作を説明する。図5は、第1の実施形態の指標算出装置の動作の一例を示すフローチャートである。

[0058] 図5に示す例では、まず、スコア算出部11が、評価対象とする系列データを入力する（ステップS101）。

[0059] 次いで、スコア算出部11は、リスケールパラメータ x_scale を取得する（ステップS102）。 x_scale は、系列データとともに入力されてもよいし、上述したような表示パラメータを基に算出してもよい。

[0060] 次いで、スコア算出部11は、取得したリスケールパラメータ x_scale に基づいて、系列データに含まれる各データに対応づけられている時間の値（ x 座標）に対してリスケールを実施する（ステップS103）。

[0061] 次いで、スコア算出部11は、リスケール済みの系列データを用いて、漸近モデルを学習する（ステップS104）。

[0062] 次いで、スコア算出部11は、リスケール済みの系列データと学習された

漸近モデルの誤差を算出する（ステップS105）。ここでは、誤差として、系列データのうち指定された各時点についての誤差関数 $loss()$ の出力の和を求める。

[0063] 最後に、スコア算出部11は、算出された誤差に基づいて、非妥当性スコアを出力する（ステップS106）。スコア算出部11は、誤差をそのまま非妥当性スコアとして出力してもよいし、例えば、算出された誤差から各時点における平均（時点平均）や、所定区間（例えば、3時点区間）における平均（区間平均）を算出し、それを非妥当性スコアとして出力してもよい。

[0064] 以上のように、本実施形態によれば、実績値を用いなくても、入力された系列データの非妥当性を判断できる客観的な指標である非妥当性スコアを出力することができる。このため、本実施形態により出力される非妥当性スコアを利用すれば、仮定の条件に基づいて得られる予測値を含む系列データの非妥当性を自動で判断できる。

[0065] 実施形態2.

次に、第2の実施形態について説明する。本実施形態では、第1の実施形態の指標算出装置10の利用例の1つとして、モデル選択機能を有する予測システムについて説明する。図6は、本実施形態の予測システムの構成例を示すブロック図である。図6に示す予測システム100は、モデル学習部101と、データ記憶部102と、予測部103と、モデル選択部104と、スコア算出部11とを備える。

[0066] なお、図6に示す例は、モデル学習部101とデータ記憶部102とを備える予測システム100に対して、モデル選択部104と予測部103とスコア算出部11とを含むモデル自動選択装置20を新たに追加した例である。ただし、予測システム100において、モデル選択部104、予測部103およびスコア算出部11は、必ずしも1つの装置に実装される構成に限定されない。

[0067] モデル学習部101は、ある検査項目の値など、所定の予測対象項目の値を予測する予測モデルの候補とされる複数のモデル候補を学習する。モデル

学習部 101 は、例えば、指定された予測対象項目の値を予測する予測モデルの候補として、説明変数の組み合わせや制約条件や各種モデルパラメータが異なる複数のモデル候補を学習してもよい。

[0068] データ記憶部 102 は、モデル学習部 101 でモデル学習に用いる学習用データや、モデル学習部 101 で学習された予測モデル候補の情報を記憶する。

[0069] また、データ記憶部 102 は、さらに、学習済みの複数のモデル候補の各々から予測値を得るために各モデル候補に入力するデータであって、各モデル候補に用いられる説明変数のデータセットである予測用データを記憶する。なお、予測用データにより得られる予測値の用途等は特に限定されない。例えば、予測用データは、予測モデル候補を検証するための検証用データであってもよいし、実際の利用者の将来の検査値等、予測システムの運用中において実際に利用される予測値を得るための予測対象データ（例えば、実際の利用者の過去の検査値と生活習慣に関する 1 つ以上の問診項目の値とを含む説明変数のデータセット等）であってもよい。

[0070] 予測部 103 は、学習済みの複数のモデル候補の各々と、データ記憶部 102 に記憶されている予測用データとを用いて経過予測を行い、モデル候補ごとに予測対象項目の系列データを生成する。

[0071] ここで、経過予測とは、少なくとも実績値を有する時点である所定の基準点から所定の予測時間単位における 2 以上の時点を含む期間であってそのうちの少なくとも 1 つの時点が予測時点とされる期間を評価対象期間として、該評価対象期間に含まれる各予測時点における予測値を得ることをいう。なお、予測時間単位は、1 年ごとに予測値を得る予測モデルであれば「1 年」とされるなど、予め予測モデルまたは予測モデル候補に定められている予測値を出力可能な標準的な時間間隔であればよい。

[0072] 例えば、予測部 103 は、学習済みの複数のモデル候補の各々にデータ記憶部 102 に記憶されている予測用データを適用して、所定の評価対象期間に含まれる各予測時点の予測値を得る。そして、予測部 103 は、モデル候

補ごとに、得られた予測値を示すデータと併せて、時間と対応づけて予測対象項目の値を示す3つ以上のデータを含む系列データを生成してもよい。

[0073] 予測部103は、経過予測を行う際、予測用データに含まれる説明変数のうちの一部の値を一定とする条件下で各モデル候補に予測用データを適用し、各予測時点の予測値を得て、得られた予測値を含む系列データを生成する。予測部103は、例えば、所定の検査値を予測するモデル候補の各々に対して、生活習慣に関する項目を一定にした予測用データを適用し、予測結果として各予測時点の予測値（検査値）を得てもよい。その場合、得られた各予測時点の予測値（検査値）を含む系列データを生成する。系列データには、予測に用いた実測値が含まれていてもよい。

[0074] スコア算出部11は、予測部103によって予測モデル候補ごとに予測結果として取得された系列データを入力とし、各系列データに対して非妥当性スコアを算出する。

[0075] モデル選択部104は、スコア算出部11によって算出された各系列データに対する非妥当性スコアに基づいて、複数の予測モデル候補の中から予測値を得る予測モデルを選択する。モデル選択部104は、例えば、非妥当性スコアが最も低いモデルを選択してもよい。なお、モデル選択部104が選択する予測モデルの数は1つに限定されず、例えば、モデル選択部104は、例えば、非妥当性スコアが低い順から所定数の予測モデルを選択してもよいし、非妥当性スコアが所定の閾値以下のモデルを全て選択することも可能である。

[0076] また、モデル選択部104は、1つの予測モデル（この場合、予測モデル候補）から複数の系列データ（例えば、複数の予測サンプルに対応した系列データ）が得られた場合には、それら複数の系列データに対する非妥当性スコアを総合してモデルを選択することも可能である。その場合において、モデル選択部104は、例えば、次の方法によりモデルを選択してもよい。

[0077] (1) モデルごとに非妥当性スコアが与えられた閾値より大きいサンプル数（系列データ）を不良サンプル数としてカウントし、不良サンプル数が少な

い順に所定数（１つでも複数でも可）のモデルを選択する。

（２）モデルごとに複数の系列データに対する非妥当性スコアの総和をとり、その値が小さい順に所定数のモデルを選択する。

（３）モデルごとに複数の系列データに対する非妥当性スコアの最大値をとり、その値が小さい順に所定数のモデルを選択する。

[0078] また、モデル選択部１０４は、予測モデルを選択する際、各系列データの非妥当性スコアと、各系列データに含まれる予測値を生成した予測モデル（本例であれば、複数の予測モデル候補）の予測精度とに基づいて、複数の予測モデル候補の中から予測値を得る予測モデルを選択することも可能である。グラフ形状に関する評価（非妥当性スコア）だけでなく、予測精度に関する評価を加えることで、予測精度と不良サンプル数のバランスが良好なモデルを選択できる。

[0079] モデル選択部１０４は、例えば、所定の検証用データ（例えば、目的変数の値が既知の説明変数の組合せからなるデータ集合）を、評価対象とされる予測モデルの各々に適用し、得られた予測値と目標値との差分を基に、予測精度（例えば、平均二乗誤差平方根（RMSE）や相関係数など）を算出してもよい。その上で、モデル選択部１０４は、例えば、予測精度が所定の閾値以上の予測モデルの中から非妥当性スコアに基づくモデル選択を行ってもよい。

[0080] 図７は、本実施形態の予測システムの動作の一例を示すフローチャートである。図７に示す例では、まず、モデル学習部１０１が、１つの予測対象項目について、複数のモデル候補を学習する（ステップＳ２０１）。

[0081] 次に、予測部１０３が、モデル学習部１０１によって学習された複数のモデル候補の各々に対して経過予測を行い、モデル候補ごとに、得られた予測値を含む系列データを生成する（ステップＳ２０２）。

[0082] 次に、スコア算出部１１が、予測部１０３によって得られたモデル候補ごとの系列データに対して、非妥当性スコアを算出する（ステップＳ２０３）。

- [0083] 最後に、モデル選択部104が、スコア算出部11によって算出された非妥当性スコアに基づいて、複数のモデル候補の中から予測値を得る予測モデルを選択する（ステップS204）。
- [0084] 以上のように、本実施形態によれば、いくつかある予測モデルの候補の中から、自動で、より妥当な経過予測が出力されるモデルを選択したり、妥当でない経過予測が出力されるモデルを除外したりすることができる。
- [0085] なお、上記の例では、モデル選択部103が少なくとも1つのモデル（予測モデル候補）を選択する例を示したが、モデル選択部103が、例えば、複数の予測モデル候補について、それらから得られた予測値を含む系列データの非妥当性スコアに基づいて、出荷の可否を判定することも可能である。その場合、例えば、モデル選択部103は、複数の予測モデル候補の各々から得られた予測値を含む系列データの非妥当性スコアに対して閾値判定を行い、所定の閾値以下であれば出荷OKとし、それ以外であれば出荷NGとしてもよい。また、モデル選択部103は、例えば、1つのモデルから複数の系列データが得られた場合には、モデルごとの不良サンプル数や非妥当性スコアの総和や最大値に対して、閾値判定を行って、所定の閾値以下であれば出荷OKとし、それ以外であれば出荷NGとしてもよい。なお、本例においても、非妥当スコアと、予測精度とに基づいて、出荷判断を行うことも可能である。その場合、モデル選択部103は、例えば、予測精度が所定の閾値以上でかつ非妥当性スコアが上記の条件を満たす場合に出荷OKと判定し、それ以外は出荷NGと判定してもよい。
- [0086] 出荷判断の結果、出荷OKであれば、モデル選択部103は、そのモデル（予測モデル候補）を出荷（外部に出力）し、出荷NGであれば、所定のアラート処理を行ってもよい。
- [0087] アラート処理としては、例えば、出荷NGとされたモデルの識別子やそのときの系列データやその非妥当性スコア等と合わせて、その旨を所定のサーバや表示装置に出力して、人手による出荷の可否判定を依頼してもよい。さらに、人手による出荷可否の判定結果を受け付けてもよい。

[0088] 実施形態 3.

次に、第3の実施形態について説明する。本実施形態では、第1の実施形態の指標算出装置10の利用例の1つとして、出荷判断機能を有する予測システムについて説明する。図8は、本実施形態の予測システムの構成例を示すブロック図である。図8に示す予測システム100は、データ記憶部102と、1つ以上の予測部103と、予測結果入力部105と、出荷判断部106と、スコア算出部11とを備える。

[0089] なお、図8に示す例は、データ記憶部102と予測部103とを備える予測システム100に対して、予測結果入力部105と出荷判断部106とスコア算出部11とを含む半自動出荷判断装置30を新たに追加した例である。ただし、予測システム100において、予測結果入力部105、出荷判断部106およびスコア算出部11は、必ずしも1つの装置に実装される構成に限定されない。

[0090] 本実施形態において、データ記憶部102は、1つ以上の予測対象項目に対応する学習済みの予測モデルの情報を記憶する。また、データ記憶部102は、それら学習済みの予測モデルから予測値を得るために各予測モデルに入力するデータであって、各予測モデルに用いられる説明変数のデータセットである予測用データ（予測サンプル）を記憶する。なお、予測用データは1つに限らず、複数であってもよい。データ記憶部102は、例えば、指定されたまたは所定の1人以上の予測対象者の各々に対応した複数の予測用データを記憶してもよい。

[0091] 予測部103はそれぞれ、1つの予測対象項目と対応づけられており、データ記憶部102に記憶されている、対応する予測対象項目の学習済みの予測モデルおよび予測用データを用いて経過予測を行い、対応する予測対象項目の系列データを生成する。

[0092] 例えば、予測部103の各々は、データ記憶部102に記憶されている対応する予測対象項目の学習済みの予測モデルを読み出し、該予測モデルに、データ記憶部102に記憶されている対応する予測対象項目の予測用データ

を適用して、対応する予測対象項目について、所定の評価対象期間に含まれる各予測時点の予測値を得る。そして、予測部103の各々は、得られた予測値を示すデータと併せて、時間と対応づけて予測対象項目の値を示す3つ以上のデータを含む系列データを生成してもよい。なお、予測部103の各々は、複数の予測用データが記憶されている場合には、予測用データごとに、予測モデルを適用して得られた予測値を示すデータを含む系列データを生成すればよい。

[0093] 予測部103の各々は、第2の実施形態と同様、経過予測を行う際、予測用データに含まれる説明変数のうちの一部の値を一定とする条件下で予測モデルに予測用データを適用し、各予測時点の予測値を得て、得られた予測値を含む系列データを生成する。予測部103の各々は、例えば、対応する所定の検査値を予測するモデル候補の各々に対して、生活習慣に関する項目を一定にした予測用データを適用し、予測結果として各予測時点の予測値（上記の検査値）を得てもよい。その場合、得られた各予測時点の予測値（検査値）を含む系列データを生成する。系列データには、予測に用いた実測値が含まれていてもよい。

[0094] 予測結果入力部105は、予測部103の各々から得られる各予測対象項目の系列データを入力する。

[0095] スコア算出部11は、予測結果入力部105より入力される各予測対象項目の系列データに対して、非妥当性スコアを算出する。

[0096] 出荷判断部106は、スコア算出部11によって算出された各予測対象項目の系列データに対する非妥当性スコアに基づいて、予測値の出荷可否の判定を行う。出荷判断部106は、例えば、1つの予測モデルから複数の系列データ（例えば、複数の予測サンプルに対応した系列データ）が得られた場合、それら全ての系列データの非妥当性スコアが所定の閾値以下であれば、各系列データに含まれる予測値を出荷OKと判定し、それ以外であれば所定のアラート処理を行ってもよい。また、出荷判断部106は、例えば、複数の予測対象項目からそれぞれ系列データが得られた場合には、それらを個別

に評価してもよいし、それらをまとめて評価（一括評価）することも可能である。一例として、出荷判断部106は、予測値が出荷される単位である出荷単位（例えば、予測対象者等）ごとに、該出荷単位における予測値を含む系列データを一括して評価してもよい。

[0097] アラート処理としては、例えば、出荷NGとされたモデルの識別子やそのときの系列データやその非妥当性スコア等と合わせてその旨を所定のサーバや表示装置に出力して、人手による出荷の可否判定を依頼してもよい。さらに、人手による出荷可否の判定結果を受け付けてもよい。

[0098] また、出荷判断部106は、そのようにして最終的に得た出荷の可否判定の結果、出荷OKであれば予測値を外部に出力してもよい。

[0099] 本実施形態においても、出荷判断部106は、予測値の出荷判断を行う際、各系列データの非妥当性スコアだけでなく、その予測値を生成した予測モデルの予測精度に基づいて、予測値の出荷可否の判定を行うことも可能である。グラフ形状に関する評価（非妥当性スコア）だけでなく、予測精度に関する評価を加えることで、予測精度と不良サンプル数のバランスが良好な予測値を出荷できる。なお、予測精度の算出方法は第2の実施形態と同様でよい。

[0100] 図9は、本実施形態の予測システムの動作の一例を示すフローチャートである。図9に示す例では、まず、予測部103の各々が、対応する予測対象項目について、学習済みの予測モデルを用いて経過予測を行い、得られた予測値を含む系列データを生成する（ステップS301）。

[0101] 次に、予測結果入力部105が、各予測対象項目について対応する予測部103の予測結果を含む系列データを入力する（ステップS302）。

[0102] 次に、スコア算出部11が、入力された各系列データに対して非妥当性スコアを算出する（ステップS303）。

[0103] 次に、出荷判断部106が、各系列データの非妥当性スコアに基づいて、得られた予測値の出荷判断を行う（ステップS304）。ここでは、出荷判断部106は、全ての系列データの非妥当性スコアが所定の閾値以下であ

るか否かによって出荷の可否を一次判定する。

- [0104] 出荷判断の結果、出荷OKであれば（ステップS305のYes）、得られた予測値を出荷（外部に出力）する（ステップS306）。
- [0105] 出荷判断の結果、出荷OKでなければ（ステップS305のNo）、所定のアラート処理を行う（ステップS307）。
- [0106] 以上のように、本実施形態によれば、予測結果を外部に出力するか否かの判断（出荷判断）において、その都度ドメインエキスパートによる目視によるチェックなどを行わなくても済むため、コストや出荷にかかる時間を削減できる。
- [0107] なお、上記の例では、出荷判断部106が、予測値に対する出荷の可否を判断する例を示したが、第2の実施形態と同様、出荷判断部106は、予測モデルに対する出荷の可否を判断することも可能である。その場合、出荷判断部106は、例えば、判断対象とされた予測モデルに所定の検証用データを適用して得られた予測値を含む1つ以上の系列データに対して算出された非妥当性スコアに基づいて、該予測モデルの出荷の可否を判定してもよい。なお、本例においても、出荷判断部106は、非妥当性スコアだけでなく、予測モデルの予測精度に基づいて、該予測モデルの出荷の可否を判定してもよい。
- [0108] また、出荷判断部106は、例えば、1つのモデルから複数のサンプルに対する予測値が得られた場合、各サンプルにおける予測値を含む系列データを総合して出荷判断を行う点については上記と同様である。なお、出荷判断部106は、例えば、複数の予測対象項目からそれぞれ系列データが得られた場合には、それらを個別に評価してもよいし、それらをまとめて評価（一括評価）することも可能である。また、出荷判断部106は、予測値や予測モデルが出荷される単位である出荷単位ごとに、該出荷単位における予測値を含む系列データを一括して評価してもよい。
- [0109] なお、上記の各実施形態では、生活習慣に関する項目を一定としたときの検査項目の予測値を含む検査値の系列データに対して、その非妥当性を評価

する方法を例示したが、予測対象や仮定の条件はこれらに限定されない。

- [0110] また、上記では、比較対象とするモデルとして漸近モデルを例示したが、比較対象とするモデルは、漸近モデル以外であってもよい。すなわち、入力された系列データに対し予め妥当とされる関数形が決まっている場合には、同様の方法で妥当とされる関数形へのフィッティングを行って比較対象とする曲線モデルを得ることにより、同様の効果を得ることができる。
- [0111] また、図10は、本発明の実施形態にかかるコンピュータの構成例を示す概略ブロック図である。コンピュータ1000は、CPU1001と、主記憶装置1002と、補助記憶装置1003と、インタフェース1004と、ディスプレイ装置1005と、入力デバイス1006とを備える。
- [0112] 上記の実施形態におけるシステム、サーバその他の装置等は、コンピュータ1000に実装されてもよい。その場合、各装置の動作は、プログラムの形式で補助記憶装置1003に記憶されていてもよい。CPU1001は、プログラムを補助記憶装置1003から読み出して主記憶装置1002に展開し、そのプログラムに従って各実施形態における所定の処理を実施する。なお、CPU1001は、プログラムに従って動作する情報処理装置の一例であり、CPU (Central Processing Unit) 以外にも、例えば、MPU (Micro Processing Unit) やMCU (Memory Control Unit) やGPU (Graphics Processing Unit)などを備えていてもよい。
- [0113] 補助記憶装置1003は、一時的でない有形の媒体の一例である。一時的でない有形の媒体の他の例として、インタフェース1004を介して接続される磁気ディスク、光磁気ディスク、CD-ROM、DVD-ROM、半導体メモリ等が挙げられる。また、このプログラムが通信回線によってコンピュータ1000に配信される場合、配信を受けたコンピュータは1000がそのプログラムを主記憶装置1002に展開し、各実施形態における所定の処理を実行してもよい。
- [0114] また、プログラムは、上記の実施形態における所定の処理の一部を実現するためのものであってもよい。さらに、プログラムは、補助記憶装置100

3に既に記憶されている他のプログラムとの組み合わせで各実施形態における所定の処理を実現する差分プログラムであってもよい。

[0115] インタフェース1004は、他の装置との間で情報の送受信を行う。また、ディスプレイ装置1005は、ユーザに情報を提示する。また、入力デバイス1006は、ユーザからの情報の入力を受け付ける。

[0116] また、実施形態における処理内容によっては、コンピュータ1000の一部の要素は省略可能である。例えば、コンピュータ1000がユーザに情報を提示しないのであれば、ディスプレイ装置1005は省略可能である。例えば、コンピュータ1000がユーザから情報入力を受け付けないのであれば、入力デバイス1006は省略可能である。

[0117] また、上記の実施形態の各構成要素の一部または全部は、汎用または専用の回路 (Circuitry)、プロセッサ等やこれらの組み合わせによって実施される。これらは単一のチップによって構成されてもよいし、バスを介して接続される複数のチップによって構成されてもよい。また、上記の実施形態の各構成要素の一部又は全部は、上述した回路等とプログラムとの組み合わせによって実現されてもよい。

[0118] 上記の実施形態の各構成要素の一部又は全部が複数の情報処理装置や回路等により実現される場合には、複数の情報処理装置や回路等は、集中配置されてもよいし、分散配置されてもよい。例えば、情報処理装置や回路等は、クライアントアンドサーバシステム、クラウドコンピューティングシステム等、各々が通信ネットワークを介して接続される形態として実現されてもよい。

[0119] 次に、本発明の概要を説明する。図11は、本発明の指標算出装置の概要を示すブロック図である。図11に示す指標算出装置60は、非妥当性スコア出力手段61を備える。

[0120] 非妥当性スコア出力手段61 (例えば、スコア算出部11) は、所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデー

タを含む系列データが入力されると、系列データの非妥当性を示す指標である非妥当性スコアであって、系列データを所定の関数形にフィッティングして得られる曲線モデルと系列データとの誤差に基づく非妥当性スコアを出力する。

[0121] 以上のような構成により、実績値を用いなくても、入力された系列データの非妥当性を判断できる客観的な指標である非妥当性スコアを出力することができる。したがって、仮定の条件に基づいて得られる予測値を含む系列データの非妥当性を自動で判断できる。

[0122] また、図12は、本発明の予測システムの概要を示すブロック図である。図12に示す予測システム600は、予測手段601と、非妥当性スコア算出手段602と、評価手段603とを備える。なお、非妥当性スコア算出手段602は、上記の指標算出装置60が備える非妥当性スコア出力手段61により実現されてもよい。

[0123] 予測手段601（例えば、予測部103）は、所定の予測対象項目について、学習済みの予測モデルを用いて所定の予測時点における予測値を得るとともに、得られた予測値を含む系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータ系列データを生成する。

[0124] 非妥当性スコア算出手段602（例えば、スコア算出部11）は、系列データの非妥当性を示す指標である非妥当性スコアであって、系列データを所定の関数形にフィッティングして得られる曲線モデルと系列データとの誤差に基づく非妥当性スコアを算出する。

[0125] 評価手段603（例えば、モデル選択部104や出荷判断部106）は、非妥当性スコア算出手段602により算出された非妥当性スコアに基づいて、系列データ、系列データに含まれる予測値または予測値を得た予測モデルを評価する。

[0126] このような構成により、予測手段601により得られた予測値を含む系列データや、該系列データに含まれる予測値や、該予測値を得た予測モデルに

ついて、妥当か否かを判定したり、いくつかある候補の中から、自動でより妥当なものを選択したり、妥当でないものを除外したりすることができる。

[0127] なお、上記の実施形態は以下の付記のようにも記載できる。

[0128] (付記1) 所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データが入力されると、前記系列データの非妥当性を示す指標である非妥当性スコアであって、前記系列データを所定の関数形にフィッティングして得られる曲線モデルと前記系列データとの誤差に基づく非妥当性スコアを出力する非妥当性スコア出力手段を備えることを特徴とする指標算出装置。

[0129] (付記2) 前記所定の関数形は、時間が無限大のときに出力値がある値に収束するとの条件を満たす所定の関数形である付記1に記載の指標算出装置。

[0130] (付記3) 前記非妥当性スコア出力手段は、前記系列データのうち所定の第1グループに属するデータを用いて前記所定の関数形へのフィッティングを行い、前記系列データのうち所定の第2グループに属するデータを用いて前記フィッティングにより得られた前記曲線モデルとの誤差を算出し、算出された前記誤差に基づいて前記非妥当性スコアを出力し、前記第1グループに属するデータと前記第2グループに属するデータとが完全一致していない付記1または付記2に記載の指標算出装置。

[0131] (付記4) 前記系列データに含まれる一部のデータを第1グループとし、前記第1グループに属しないデータまたは前記系列データの全てのデータを第2グループとする付記3に記載の指標算出装置。

[0132] (付記5) 前記非妥当性スコア出力手段は、前記所定の関数形へのフィッティングを行う前に、前記系列データに含まれる各データに対応づけられている時間の値を、前記系列データの表示スケールに合わせて変換する付記1から付記4のうちのいずれか1項に記載の指標算出装置。

[0133] (付記6) 所定の予測対象項目について、学習済みの予測モデルを用いて

所定の予測時点における予測値を得るとともに、得られた前記予測値を含む系列データであって時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データを生成する予測手段と、前記系列データの非妥当性を示す指標である非妥当性スコアであって、前記系列データを所定の関数形にフィッティングして得られる曲線モデルと前記系列データとの誤差に基づく非妥当性スコアを算出する非妥当性スコア算出手段と、前記系列データ、前記系列データに含まれる予測値または前記予測値を得た予測モデルに対して、前記非妥当性スコアに基づく評価を行う評価手段とを備えることを特徴とする予測システム。

[0134] (付記7) 前記評価手段は、前記系列データ、前記系列データに含まれる予測値または前記予測値を得た予測モデルに対して、前記非妥当性スコアと、所定の検証用データを用いて算出される前記予測モデルの予測精度とに基づいて評価を行う付記6に記載の予測システム。

[0135] (付記8) 前記所定の予測対象項目について、複数のモデル候補を学習するモデル学習手段を備え、前記予測手段は、前記複数のモデル候補の各々を用いて前記予測値を得るとともに、前記複数のモデル候補ごとに、得られた前記予測値を含む系列データを生成し、前記非妥当性スコア算出手段は、前記複数のモデル候補ごとの前記系列データに対して前記非妥当性スコアを算出し、前記評価手段は、前記複数のモデル候補に対して前記評価を行い、その評価結果に基づいて前記複数のモデル候補の中から前記予測対象項目の予測値を得るモデルを選択する付記6または付記7に記載の予測システム。

[0136] (付記9) 前記予測手段は、1つ以上の予測対象項目の各々に対応して1つ以上設けられ、前記非妥当性スコア出力手段は、前記1つ以上の予測手段から得られる前記予測対象項目ごとの系列データに対して前記非妥当性スコアを算出し、前記評価手段は、前記予測対象項目ごとの系列データに対して前記評価を行い、その評価結果に基づいて前記予測値の出荷判断を行う付記6または付記7に記載の予測システム。

[0137] (付記10) 情報処理装置が、所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データが入力されると、前記系列データの非妥当性を示す指標である非妥当性スコアであって、前記系列データを所定の関数形にフィッティングして得られる曲線モデルと前記系列データとの誤差に基づく非妥当性スコアを算出し、前記非妥当性スコアに基づいて、前記系列データまたは前記系列データに含まれる予測値もしくは前記予測値を得た予測モデルを評価することを特徴とする経過予測評価方法。

[0138] (付記11) コンピュータに、所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データが入力されると、前記系列データの非妥当性を示す指標である非妥当性スコアであって、前記系列データを所定の関数形にフィッティングして得られる曲線モデルと前記系列データとの誤差に基づく非妥当性スコアを算出する処理、および前記非妥当性スコアに基づいて、前記系列データまたは前記系列データに含まれる予測値もしくは前記予測値を得た予測モデルを評価する処理を実行させるための経過予測評価プログラム。

[0139] 以上、本実施形態および実施例を参照して本願発明を説明したが、本願発明は上記実施形態および実施例に限定されるものではない。本願発明の構成や詳細には、本願発明の範囲内で当業者が理解し得る様々な変更をすることができる。

[0140] この出願は、2018年3月30日に提出された日本特許出願2018-067622を基礎とする優先権を主張し、その開示の全てをここに取り込む。

産業上の利用の可能性

[0141] 本発明は、仮定の条件に基づいて得られる予測値を含む系列データだけでなく、時間と対応づけて予測対象項目の値を示す3つ以上のデータを含む系

列データについて妥当な関数形が与えられるものであれば、好適に適用可能である。

符号の説明

- [0142] 10 指標算出装置
- 11 スコア算出部
- 20 モデル自動選択装置
- 30 半自動出荷判断装置
- 100 予測システム
- 101 モデル学習部
- 102 データ記憶部
- 103 予測部
- 104 モデル選択部
- 105 予測結果入力部
- 106 出荷判断部
- 1000 コンピュータ
- 1001 CPU
- 1002 主記憶装置
- 1003 補助記憶装置
- 1004 インタフェース
- 1005 ディスプレイ装置
- 1006 入力デバイス
- 60 指標算出装置
- 61 非妥当性スコア出力手段
- 600 予測システム
- 601 予測手段
- 602 非妥当性スコア算出手段
- 603 評価手段

請求の範囲

- [請求項1] 所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データが入力されると、前記系列データの非妥当性を示す指標である非妥当性スコアであって、前記系列データを所定の関数形にフィッティングして得られる曲線モデルと前記系列データとの誤差に基づく非妥当性スコアを出力する非妥当性スコア出力手段を備える
- ことを特徴とする指標算出装置。
- [請求項2] 前記所定の関数形は、時間が無限大のときに出力値がある値に収束するとの条件を満たす所定の関数形である
- 請求項1に記載の指標算出装置。
- [請求項3] 前記非妥当性スコア出力手段は、前記系列データのうち所定の第1グループに属するデータを用いて前記所定の関数形へのフィッティングを行い、前記系列データのうち所定の第2グループに属するデータを用いて前記フィッティングにより得られた前記曲線モデルとの誤差を算出し、算出された前記誤差に基づいて前記非妥当性スコアを出力し、
- 前記第1グループに属するデータと前記第2グループに属するデータとが完全一致していない
- 請求項1または請求項2に記載の指標算出装置。
- [請求項4] 前記系列データに含まれる一部のデータを第1グループとし、前記第1グループに属しないデータまたは前記系列データの全てのデータを第2グループとする
- 請求項3に記載の指標算出装置。
- [請求項5] 前記非妥当性スコア出力手段は、前記所定の関数形へのフィッティングを行う前に、前記系列データに含まれる各データに対応づけられている時間の値を、前記系列データの表示スケールに合わせて変換す

る

請求項 1 から請求項 4 のうちのいずれか 1 項に記載の指標算出装置

。

[請求項6]

所定の予測対象項目について、学習済みの予測モデルを用いて所定の予測時点における予測値を得るとともに、得られた前記予測値を含む系列データであって時間と対応づけて予測対象項目の値を示す 3 つ以上のデータであってそのうちの少なくとも 1 つが予測値を示すデータを含む系列データを生成する予測手段と、

前記系列データの非妥当性を示す指標である非妥当性スコアであって、前記系列データを所定の関数形にフィッティングして得られる曲線モデルと前記系列データとの誤差に基づく非妥当性スコアを算出する非妥当性スコア算出手段と、

前記系列データ、前記系列データに含まれる予測値または前記予測値を得た予測モデルに対して、前記非妥当性スコアに基づく評価を行う評価手段とを備える

ことを特徴とする予測システム。

[請求項7]

前記評価手段は、前記系列データ、前記系列データに含まれる予測値または前記予測値を得た予測モデルに対して、前記非妥当性スコアと、所定の検証用データを用いて算出される前記予測モデルの予測精度とに基づいて評価を行う

請求項 6 に記載の予測システム。

[請求項8]

前記所定の予測対象項目について、複数のモデル候補を学習するモデル学習手段を備え、

前記予測手段は、前記複数のモデル候補の各々を用いて前記予測値を得るとともに、前記複数のモデル候補ごとに、得られた前記予測値を含む系列データを生成し、

前記非妥当性スコア算出手段は、前記複数のモデル候補ごとの前記系列データに対して前記非妥当性スコアを算出し、

前記評価手段は、前記複数のモデル候補に対して前記評価を行い、その評価結果に基づいて前記複数のモデル候補の中から前記予測対象項目の予測値を得るモデルを選択する

請求項6または請求項7に記載の予測システム。

[請求項9]

前記予測手段は、1つ以上の予測対象項目の各々に対応して1つ以上設けられ、

前記非妥当性スコア出力手段は、前記1つ以上の予測手段から得られる前記予測対象項目ごとの系列データに対して前記非妥当性スコアを算出し、

前記評価手段は、前記予測対象項目ごとの系列データに対して前記評価を行い、その評価結果に基づいて前記予測値の出荷判断を行う

請求項6または請求項7に記載の予測システム。

[請求項10]

情報処理装置が、

所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データが入力されると、前記系列データの非妥当性を示す指標である非妥当性スコアであって、前記系列データを所定の関数形にフィッティングして得られる曲線モデルと前記系列データとの誤差に基づく非妥当性スコアを算出し、

前記非妥当性スコアに基づいて、前記系列データまたは前記系列データに含まれる予測値もしくは前記予測値を得た予測モデルを評価する

ことを特徴とする経過予測評価方法。

[請求項11]

コンピュータに、

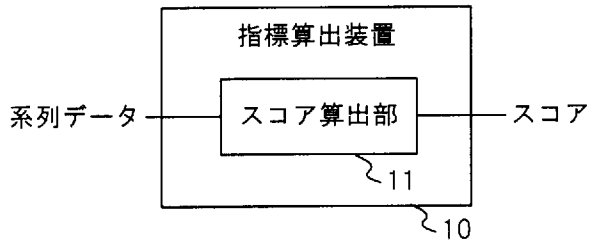
所定の予測対象項目の系列データであって、時間と対応づけて予測対象項目の値を示す3つ以上のデータであってそのうちの少なくとも1つが予測値を示すデータを含む系列データが入力されると、前記系列データの非妥当性を示す指標である非妥当性スコアであって、前記

系列データを所定の関数形にフィッティングして得られる曲線モデルと前記系列データとの誤差に基づく非妥当性スコアを算出する処理、および

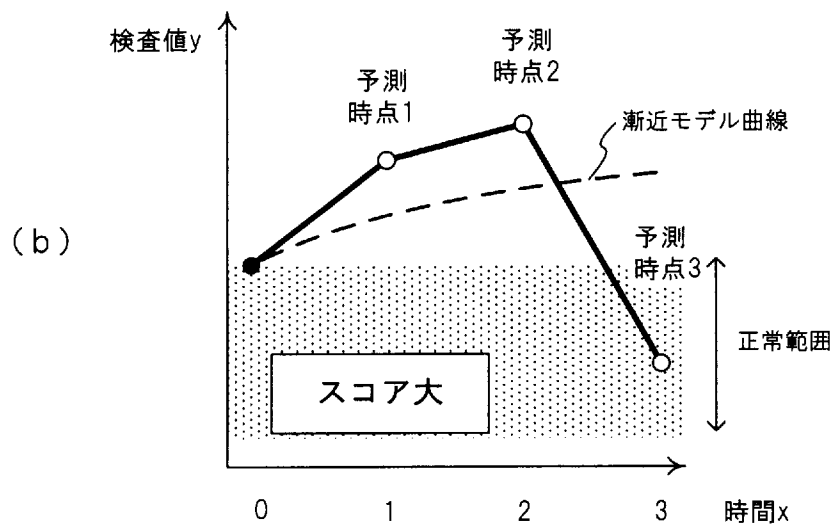
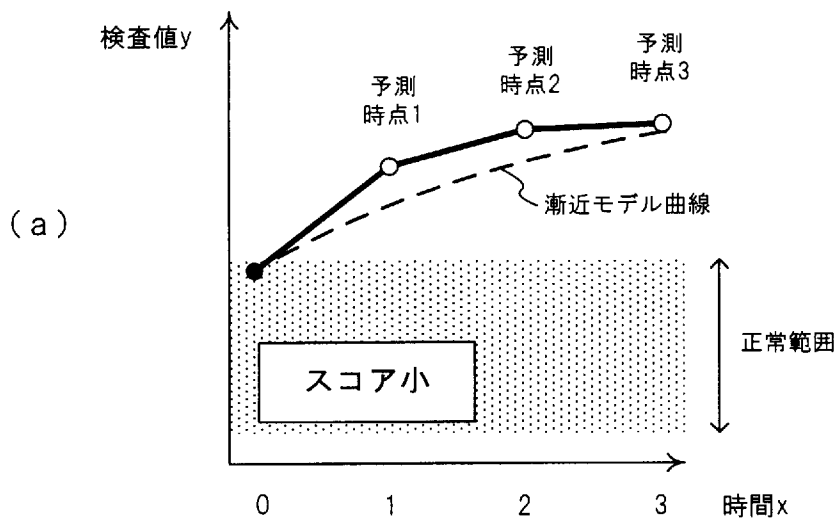
前記非妥当性スコアに基づいて、前記系列データまたは前記系列データに含まれる予測値もしくは前記予測値を得た予測モデルを評価する処理

を実行させるための経過予測評価プログラム。

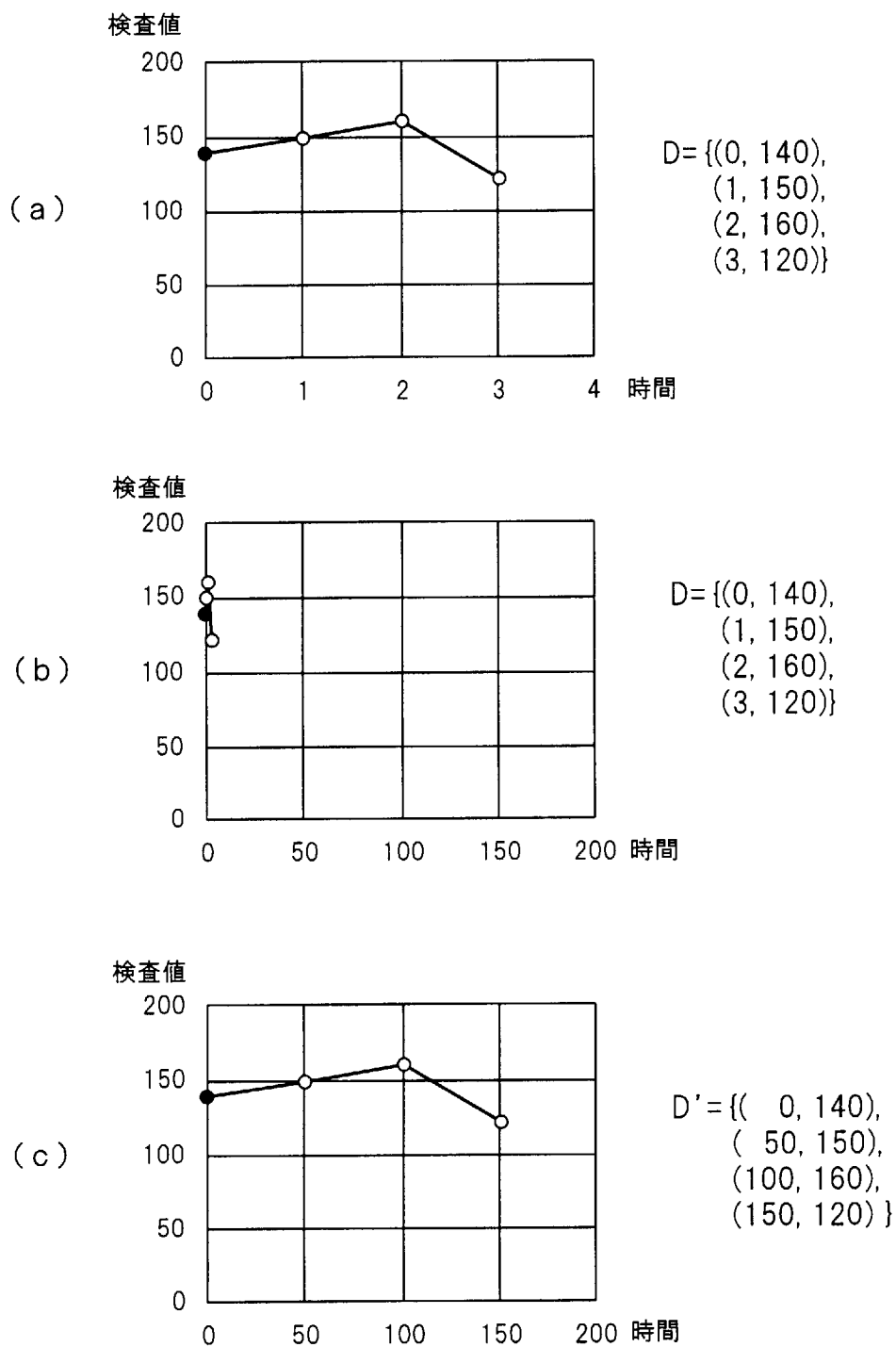
[図1]



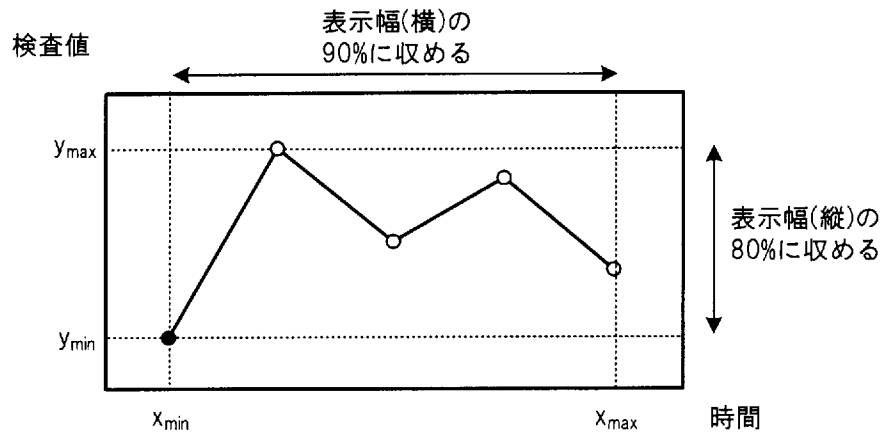
[図2]



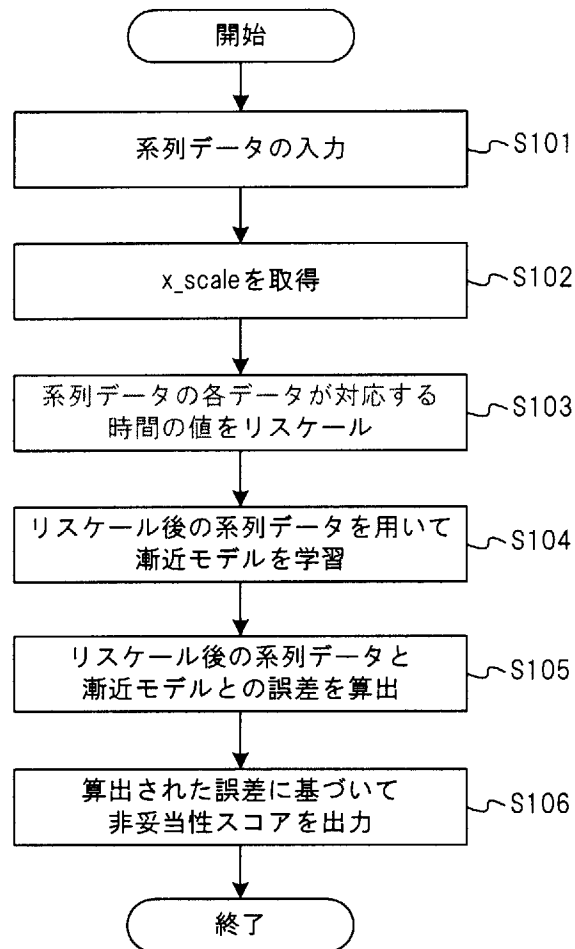
[図3]



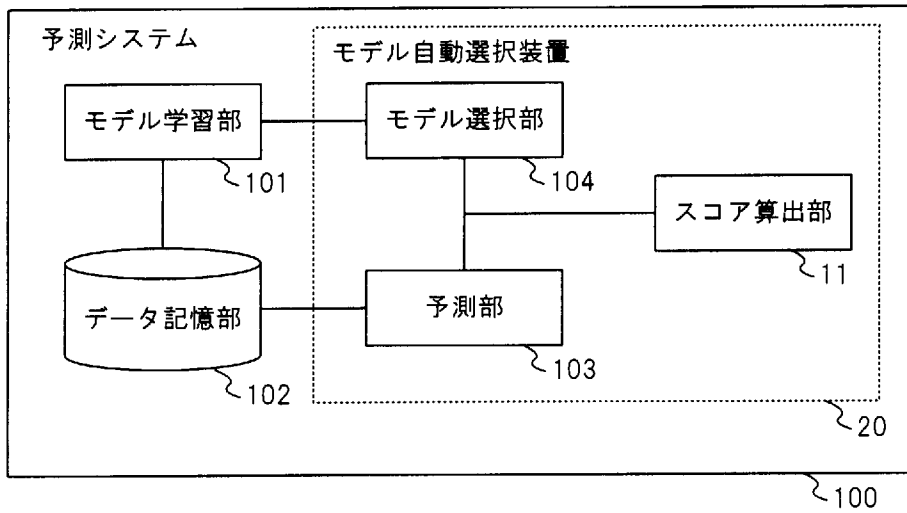
[図4]



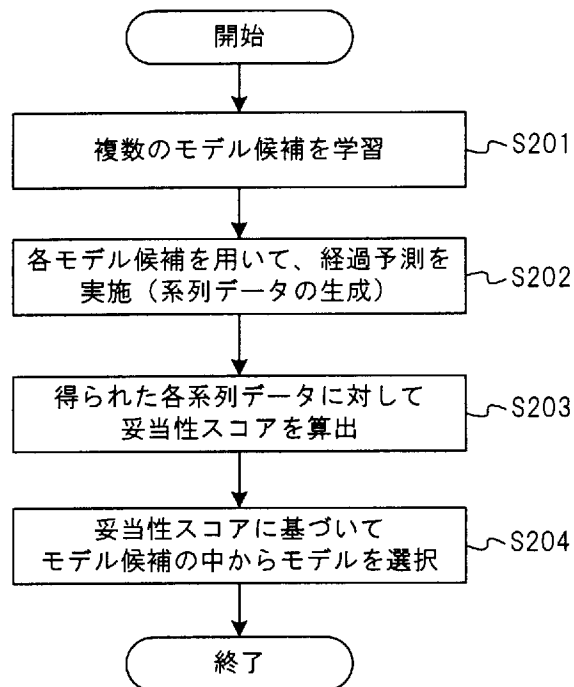
[図5]



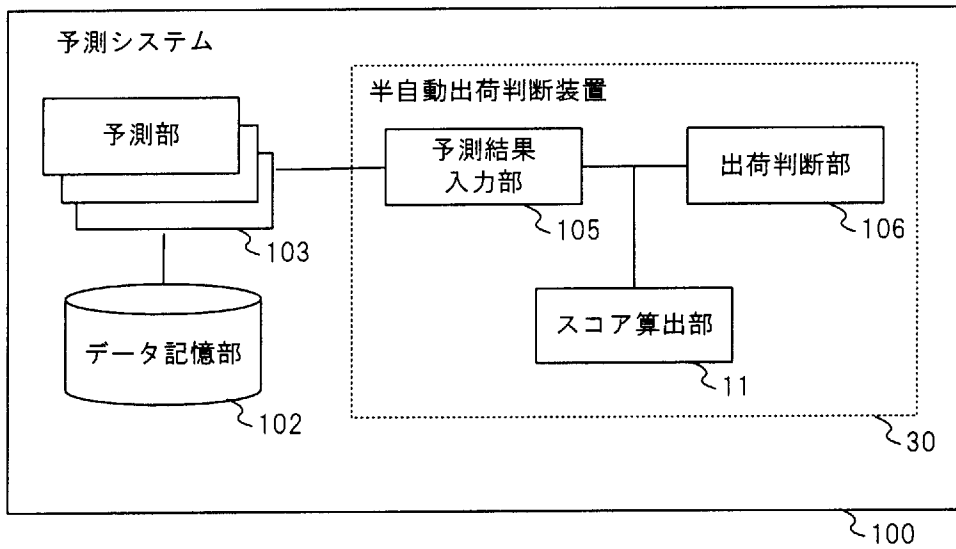
[図6]



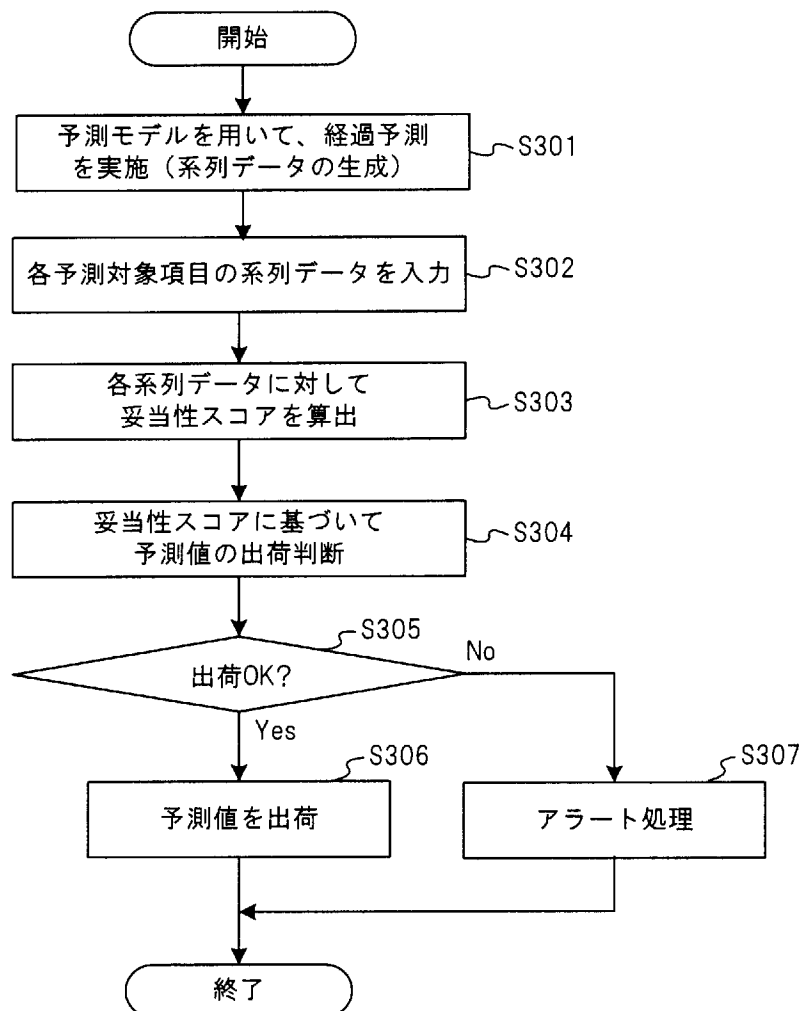
[図7]



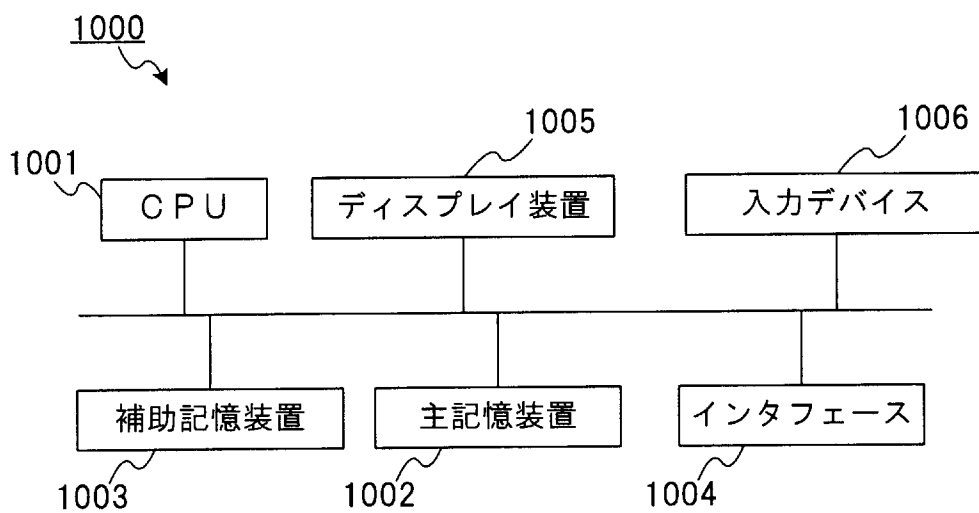
[図8]



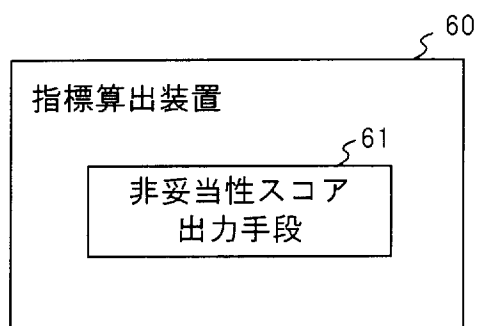
[図9]



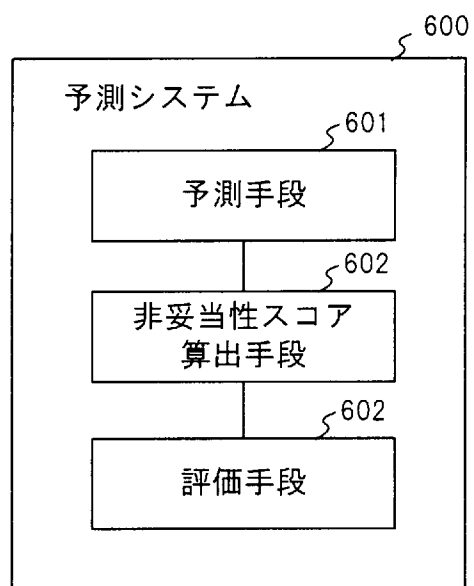
[図10]



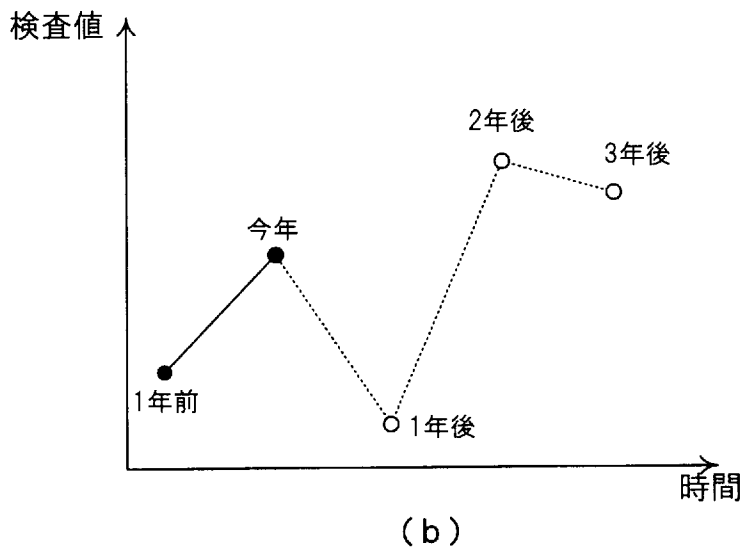
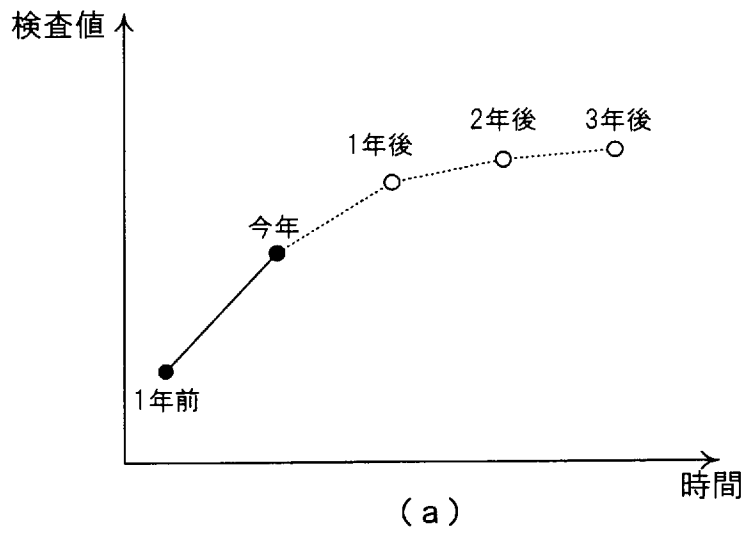
[図11]



[図12]



[図13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2018/043910

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06Q10/04 (2012.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. G06Q10/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	JP 2011-176984 A (TOSHIBA CORPORATION) 08 September 2011, paragraphs [0069], [0070] (Family: none)	1-7, 9-11 8
Y A	JP 2016-201869 A (LEMONGAS CO., LTD.) 01 December 2016, paragraphs [0063], [0073] (Family: none)	1-7, 9-11 8
Y A	JP 2005-63208 A (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) 10 March 2005, paragraph [0003], fig. 5 (Family: none)	2-5 1, 6-11

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
 “E” earlier application or patent but published on or after the international filing date
 “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 “O” document referring to an oral disclosure, use, exhibition or other means
 “P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
 “&” document member of the same patent family

Date of the actual completion of the international search
06.02.2019

Date of mailing of the international search report
19.02.2019

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2018/043910

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	WO 2005/031278 A1 (NOZZLE NETWORK CO., LTD.) 07 April 2005, paragraphs [0079]-[0081] (Family: none)	5 1-4, 6-11
A	JP 2017-49677 A (FUJITSU LIMITED) 09 March 2017, entire text, all drawings & US 2017/0061329 A1, entire text, all drawings	1-11

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G06Q10/04(2012.01)i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl. G06Q10/04

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y A	JP 2011-176984 A (株式会社東芝) 2011.09.08, 段落[0069]-[0070] (ファミリーなし)	1-7, 9-11 8
Y A	JP 2016-201869 A (レモンガス株式会社) 2016.12.01, 段落[0063][0073] (ファミリーなし)	1-7, 9-11 8
Y A	JP 2005-63208 A (日本電信電話株式会社) 2005.03.10, 段落[0003][図5] (ファミリーなし)	2-5 1, 6-11

☑ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

- 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
- 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
- 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
- 「O」 口頭による開示、使用、展示等に言及する文献
- 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

- 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
- 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
- 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
- 「&」 同一パテントファミリー文献

国際調査を完了した日

06.02.2019

国際調査報告の発送日

19.02.2019

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
郵便番号 100-8915
東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

阿部 潤

電話番号 03-3581-1101 内線 3562

5L

4173

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y A	WO 2005/031278 A1 (ノズルネットワーク株式会社) 2005. 04. 07, 段落[0079]-[0081] (ファミリーなし)	5 1-4, 6-11
A	JP 2017-49677 A (富士通株式会社) 2017. 03. 09, 全文、全図 & US 2017/0061329 A1 全文、全図	1-11