



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2011-0097528
(43) 공개일자 2011년08월31일

(51) Int. Cl.

G06Q 50/00 (2006.01)

(21) 출원번호 10-2010-0017411

(22) 출원일자 2010년02월25일

심사청구일자 2010년02월25일

(71) 출원인

고려대학교 산학협력단

서울 성북구 안암동5가 1

(72) 발명자

이재원

서울특별시 서초구 반포동 반포@ 84-401

김호

서울특별시 강남구 도곡동 아이파크3차 104-301

(뒷면에 계속)

(74) 대리인

특허법인충현

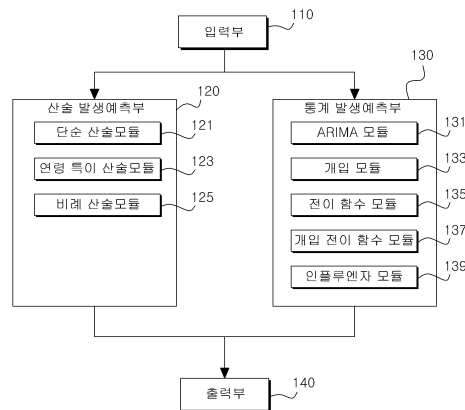
전체 청구항 수 : 총 11 항

(54) 백신 예방가능한 질병의 발생 예측 시스템

(57) 요약

데이터가 입력되는 입력장치, 입력된 데이터로부터 질병 발생에 관한 연산을 수행하는 컴퓨터 및 연산된 질병 발생에 관한 수치를 출력하는 출력장치를 포함하는 시스템을 개시한다. 상기 시스템은 예방접종 정책에 대한 정보 및 상기 정보 관련된 모수를 입력받아 미분방정식을 이용하여 질병 발생수의 예측값을 출력하는 산술 발생예측부 및 과거 질병 발생수를 입력받아 시계열적으로 질병 발생수의 예측값을 출력하는 통계 발생예측부를 포함한다. 통계 또는 수학 전문 프로그램에 대한 지식이 없어도 질병 발생을 예측할 수 있으며, 질병의 패턴 등을 손쉽게 파악할 수 있다.

대표도 - 도1



(72) 발명자

허만희

대전광역시 유성구 반석동 반석마을 8단지 802동 301호

김성주

경기도 부천시 원미구 심곡1동 372-14

유아미

서울특별시 성북구 안암동5가 147-9 201호

최병엽

서울특별시 성북구 안암동3가 맥시칸 빌라 102호

홍현숙

서울특별시 동대문구 이문동 292-120번지 2층

이 발명을 지원한 국가연구개발사업

과제고유번호 2007E00052001

부처명 질병관리본부

연구관리전문기관

연구사업명 학술연구용역사업

연구과제명 VPD 발생예측 알고리즘 연구

기여율

주관기관 고려대학교 산학협력단

연구기간 2007년 10월 31일 ~ 2008년 06월 30일

특허청구의 범위

청구항 1

데이터가 입력되는 입력장치, 입력된 데이터로부터 질병 발생에 관한 연산을 수행하는 컴퓨터 및 연산된 질병 발생에 관한 수치를 출력하는 출력장치를 포함하는 시스템에 있어서,

예방접종 정책에 대한 정보 및 상기 정보 관련된 모수를 입력받아 미분방정식을 이용하여 질병 발생수의 예측값을 출력하는 산술 발생예측부 및

과거 질병 발생수를 입력받아 시계열적으로 질병 발생수의 예측값을 출력하는 통계 발생예측부를 포함하는 질병 발생 예측 시스템.

청구항 2

제 1 항에 있어서,

상기 질병은 백신 예방가능한 질병인 것을 특징으로 하는 질병 발생 예측 시스템.

청구항 3

제 2 항에 있어서,

상기 질병은 A형간염, B형간염, 홍역, 유행성이하선염, 풍진, 파상풍, 백일해, 일본뇌염, 인플루엔자, 수두, 소아마비 중 적어도 하나인 것을 특징으로 하는 질병 발생 예측 시스템.

청구항 4

제 1 항에 있어서,

상기 산술 발생예측부는

감수성(susceptible)집단, 잠재(infected)집단, 감염(infectious)집단 및 회복(recovery)집단에 관한 초기값, 시뮬레이션 기간(time horizon), 전파력(transmission rate), 잠재기(Latency period) 및 회복기(Recovery period)에 대한 정보를 입력받아 미분방정식을 이용하여 감수성기, 잠재기, 감염기 및 회복기에서의 질병 발생수의 예측값을 출력하는 단순 산술모델 및

전염병학(epidemiology)에 관한 정보, 회복기, 신생아수, 사망율 및 백신접종에 관한 정보를 입력받아 미분방정식을 이용하여 연령군에 따른 질병 발생수의 예측값을 출력하는 연령-특이 산술모델을 포함하는 것을 특징으로 하는 질병 발생 예측 시스템.

청구항 5

제 4 항에 있어서,

상기 전염병학(epidemiology)에 관한 정보는 그룹에 따른 연령 증가율(aging rate) 및 잠재기(Latency period) 중 적어도 하나인 것을 특징으로 하는 질병 발생 예측 시스템.

청구항 6

제 4 항에 있어서,

상기 백신접종에 관한 정보는 시뮬레이션 기간, 백신접종이 시행되는 시기, 백신효능, 접종율 및 백신을 시행할 그룹 중 적어도 하나인 것을 특징으로 하는 질병 발생 예측 시스템.

청구항 7

제 4 항에 있어서,

상기 산술 발생예측부는

전염병학(epidemiology)에 관한 정보, 회복기, 신생아수, 사망율, 백신접종에 관한 정보 및 출생률을 입력받아 감수성(susceptible)집단의 연령군에서의 질병 발생수의 예측값을 출력하는 비례 산술모듈을 더 포함하는 것을 특징으로 하는 질병 발생 예측 시스템.

청구항 8

제 1 항에 있어서,

상기 통계 발생예측부는

시간 및 질병 발생수를 입력받아 현재 관측값과 과거의 관측값 간에 상관성이 존재하는 시계열 자료로부터 시간에 따른 질병 발생수 및 예측값을 출력하는 ARIMA 모듈;

상기 시간, 질병 발생수 및 개입(intervention)이 발생했을 경우 개입 변수를 입력받아 시간에 따른 질병 발생수 및 예측값을 출력하는 개입 모듈; 및

상기 시간, 질병 발생수 및 전이 함수 변수를 입력받아 시간에 따른 질병 발생수 및 예측값을 출력하는 전이 함수 모듈을 포함하는 것을 특징으로 하는 질병 발생 예측 시스템.

청구항 9

제 8 항에 있어서,

상기 개입 모듈은 펄스 함수 및 계단 함수 중 적어도 하나를 이용하는 것을 특징으로 하는 질병 발생 예측 시스템.

청구항 10

제 1 항에 있어서,

R 언어로 구현된 모델메이커(ModelMaker) 및 SAS 프로그램의 기능을 C 계열의 언어를 이용하여 인터페이스화된 것을 특징으로 하는 질병 발생 예측 시스템.

청구항 11

제 10 항에 있어서,

상기 C 계열의 언어는 C#(sharp)인 것을 특징으로 하는 질병 발생 예측 시스템.

명세서

기술분야

[0001] 본 발명은 질병 발생 예측 시스템에 관한 것으로, 특히 VPD(Vaccine Preventable Disease)의 발생 예측 시스템에 관한 것이다.

배경기술

[0002] VPD(Vaccine Preventable Disease)는 예방접종으로 예방이 가능한 질병들, 예를 들어 A형간염, B형간염, 홍역, 유행성이하선염, 풍진, 파상풍, 백일해, 일본뇌염, 인플루엔자, 수두, 소아마비 등을 의미한다. 예방접종을 하지 않는 경우 백신으로 예방 가능한 질병의 대유행(outbreak)을 초래할 수 있으며 특히, 영유아의 경우 그 피해가 더 클 수 있다. 하지만 VPD의 발생을 예측하고 조기에 유행을 억제하는 정책을 수립하여 향후 해당 VPD를 퇴치(elimination)하거나 박멸(eradication)을 하는 것은 국가 정책적으로 매우 중요하다.

[0003] 그러나, 국내에서는 VPD 발생 예측 프로그램에 관한 연구가 매우 미진한 실정이므로 국외 연구 동향의 실태 조사를 통하여 얻어지는 지식들이 귀중한 데이터로 이용된다. 미국의 질병관리본부(Centers for Disease Control and Prevention)에서 제공하는 소프트웨어인 'Comprehensive Clinic Assessment Software Application (CoCASA)'라는 프로그램 또는 미국의 트루만주 대학이 개발한 196 - SIR - Model이라는 프로그램 등이 있다. "Epidemic simulator for Zoo4 project"라는 프로그램은 인플루엔자와 홍역에 관한 전염병 예측에 관하여 시뮬

레이션을 수행하는 프로그램이다. 하지만 국내에서는 프로그램에 관한 전문적인 지식이 없는 사용자가 손쉽게 질병발생을 예측할 수 있는 알고리즘이나 시스템은 없는 실정이다.

발명의 내용

해결하려는 과제

[0004] 본 발명이 이루고자 하는 기술적 과제는 통계 또는 수학 전문 프로그램에 대한 지식이 없어도 질병 발생을 예측할 수 있으며, 감염병의 패턴 등을 손쉽게 파악할 수 있는 질병 발생 예측 시스템을 제공하는 데 있다.

과제의 해결 수단

[0005] 본 발명의 일 양태에 따르면 데이터가 입력되는 입력장치, 입력된 데이터로부터 질병 발생에 관한 연산을 수행하는 컴퓨터 및 연산된 질병 발생에 개대한 정보 및 상기 정보 관련된 모수를 입력받아 미분방정식을 이용하여 질병 발생수의 예측값을 출력하는 산술 발생예측부 및 과거 질병 발생수를 입력받아 시계열적으로 질병 발생수의 예측값을 출력하는 통계 발생예측부를 포함하는 질병 발생 예측 시스템을 제공한다.

[0006] 상기 질병은 백신 예방가능한 질병일 수 있다.

[0007] 상기 질병은 A형간염, B형간염, 홍역, 유행성이하선염, 풍진, 파상풍, 백일해, 일본뇌염, 인플루엔자, 수두, 소아마비 중 적어도 하나일 수 있다.

[0008] 상기 산술 발생예측부는 감수성(susceptible)집단, 잠재(infected)집단, 감염(infectious)집단 및 회복(recovery)집단에 관한 초기값, 시뮬레이션 기간(time horizon), 전파력(transmission rate), 잠재기(Latency period) 및 회복기(Recovery period)에 대한 정보를 입력받아 미분방정식을 이용하여 감수성기, 잠재기, 감염기 및 회복기에서의 질병 발생수의 예측값을 출력하는 단순 산술모듈 및 전염병학(epidemiology)에 관한 정보, 회복기, 신생아수, 사망율 및 백신접종에 관한 정보를 입력받아 미분방정식을 이용하여 연령군에 따른 질병 발생수의 예측값을 출력하는 연령-특이 산술모듈을 포함할 수 있다.

[0009] 상기 전염병학(epidemiology)에 관한 정보는 그룹에 따른 연령 증가율(aging rate) 및 잠재기(Latency period) 중 적어도 하나일 수 있다.

[0010] 상기 백신접종에 관한 정보는 시뮬레이션 기간, 백신접종이 시행되는 시기, 백신효능, 접종율 및 백신을 시행할 그룹 중 적어도 하나일 수 있다.

[0011] 상기 산술 발생예측부는 전염병학(epidemiology)에 관한 정보, 회복기, 신생아수, 사망율, 백신접종에 관한 정보 및 출생률을 입력받아 감수성(susceptible)집단의 연령군에서의 질병 발생수의 예측값을 출력하는 비례 산술모듈을 더 포함할 수 있다.

[0012] 상기 통계 발생예측부는 시간 및 질병 발생수를 입력받아 현재 관측값과 과거의 관측값 간에 상관관계 존재하는 시계열 자료로부터 시간에 따른 질병 발생수 및 예측값을 출력하는 ARIMA 모듈, 상기 시간, 질병 발생수 및 개입(intervention)이 발생했을 경우 개입 변수를 입력받아 시간에 따른 질병 발생수 및 예측값을 출력하는 개입모듈, 및 상기 시간, 질병 발생수 및 전이 함수 변수를 입력받아 시간에 따른 질병 발생수 및 예측값을 출력하는 전이 함수 모듈을 포함할 수 있다.

[0013] 상기 개입 모듈은 펄스 함수 및 계단 함수 중 적어도 하나를 이용할 수 있다.

[0014] 상기 질병 발생 예측 시스템은 R 언어로 구현된 모델메이커(ModelMaker) 및 SAS 프로그램의 기능을 C 계열의 언어를 이용하여 인터페이스화될 수 있다.

[0015] 상기 C 계열의 언어는 C#(sharp)일 수 있다.

발명의 효과

[0016] 본 실시예에 따른 질병 발생 예측 시스템은 통계 또는 수학 전문 프로그램에 대한 지식이 없어도 질병 발생을 예측할 수 있으며, 질병의 패턴 등을 손쉽게 파악할 수 있고, 사용자 친화적 인터페이스(User-friendly interface)를 갖추고 있다. 종래 개별 프로그램으로 사용되던 SAS, R 및 ModelMaker 등이 하나의 시스템으로 통합된다.

도면의 간단한 설명

- [0017] 도 1은 본 실시예에 따른 질병 발생 예측 시스템의 개략적인 구성을 나타내는 구성도이다.
- 도 2는 통계 발생예측부의 입력 변수의 일례를 보여주는 화면이다.
- 도 3은 본 실시예에 따른 질병 발생 예측 시스템의 개발 과정을 설명하는 도면이다.
- 도 4는 본 실시예에 따른 질병 발생 예측 시스템의 초기 입력인터페이스 화면이다.
- 도 5는 본 실시예에 따른 연령-특이 산술모듈의 변수 설정 화면이다.
- 도 6은 본 실시예에 따른 연령-특이 산술모듈의 출력화면이다.
- 도 7은 본 실시예에 따른 개입 전이 함수 모듈의 변수 설정 화면이다.
- 도 8은 본 실시예에 따른 개입 전이 함수 모듈의 출력화면이다.

발명을 실시하기 위한 구체적인 내용

- [0018] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0019] 제2, 제1 등과 같이 서수를 포함하는 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되지는 않는다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제2 구성요소는 제1 구성요소로 명명될 수 있고, 유사하게 제1 구성요소도 제2 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0020] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.
- [0021] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0022] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0023] 이하, 첨부된 도면을 참조하여 실시예를 상세히 설명하되, 도면 부호에 관계없이 동일하거나 대응하는 구성 요소는 동일한 참조 번호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다.
- [0024] 도 1은 본 실시예에 따른 질병 발생 예측 시스템의 개략적인 구성을 나타내는 구성도이다. 본 실시예에 따른 질병 발생 예측 시스템은 데이터 입력장치를 통해 수신되는 입력신호에 따라 데이터를 연산하여 질병의 발생을 예측하여 출력장치를 통해 디스플레이하는 시스템으로서 데이터의 연산을 수행하는 컴퓨터 및 과거 질병 발생에 관한 데이터베이스를 포함하여 이루어지며, 입력부(110), 산술 발생예측부(120), 통계 발생예측부(130) 및 출력부(140)를 포함한다.
- [0025] 입력부(110)는 각 모듈에 따른 설정 변수를 입력받는다. 산술 발생예측부(120)는 단순 산술모듈(121), 연령-특이 산술모듈(123) 및 비례 산술모듈(125)을 포함하고 예방접종 정책에 대한 정보나 관련 모수를 이용하여 질병 발생수를 예측하여 출력하며, 통계 발생예측부(130)는 ARIMA 모듈(131), 개입 모듈(133), 전이 함수 모듈(135), 개입 전이 함수 모듈(137) 및 인플루엔자 모듈(139)을 포함하고 과거 질병 발생수를 이용하여 시계열적으로 질

병 발생수를 예측하여 출력한다.

[0026] 단순 산술모듈(121)은 감수성기, 잠재기, 감염기 및 회복기 단계별로 집단이 바이러스에 노출되면서 감염에 의해서 집단 내 면역수준이 변화하는 양상을 단위 시간에 따른 평균변화속도로 나타낸 미분방정식을 이용하여 질병 발생수를 예측하여 출력한다. 감수성(susceptible)집단, 잠재(infected)집단, 감염(infectious)집단 및 회복(recovery)집단에 관한 초기값, 시물레이션 기간(time horizon), 전파력(transmission rate), 잠재기(Latency period) 및 회복기(Recovery period)에 대한 정보를 입력받아 다양한 변수를 포함하는 미분방정식을 이용하여 감수성기, 잠재기, 감염기 및 회복기에서의 질병 발생수를 예측하여 출력한다. 여기서, 질병에 따라 상기 집단 중 일부가 없을 수 있으며 다른 일부가 더 추가될 수도 있음은 당업자에게 자명할 것이다.

[0027] 여기서, 감수성(susceptible)집단은 전염병에 감염될 수 있는 집단을 나타내고, 예를 들어 100을 입력하는 경우 전염병에 감염될 수 있는 집단이 100명임을 나타낸다. 잠재(infected)집단은 감염은 되었으나 감수성집단에게 감염을 시킬 수 있는 단계는 아닌 집단을 나타낸다. 감염(infectious)집단은 감수성집단에게 감염시킬 수 있는 집단을 나타낸다. 회복(recovery)집단은 감염의 결과로 면역을 획득한 집단을 나타낸다. 시물레이션 기간(days)은 질병 발생을 예측하는 기간으로 예를 들어, 365일을 입력한다면 365일간 질병 발생을 예측하게 된다. 전파력(transmission rate)은 감염된 사람이 감수성집단과 접촉하여 실제 감염시킬 수 있는 확률을 말한다. 잠재기(Latency period; days)는 숙주가 감염되어 다른 사람을 감염시킬 수 있는 기간을 말한다. 짧을수록 질병의 전파가 가속화된다. 회복기(Recovery period; days)는 병원체를 배출하기 시작한 시점에서 배출을 종료한 시점까지 걸리는 기간을 말한다.

[0028] 연령-특이 산술모듈(123)은 전염병학(epidemiology)에 관한 정보, 예를 들어, 그룹에 따른 연령 증가율(aging rate)와 잠재기, 회복기, 신생아수, 사망율, 백신접종에 관한 정보를 입력받아 다양한 변수를 포함하는 미분방정식을 이용하여 연령군에 따른 질병 발생수를 예측한다. 연령-특이 산술모듈(123)에서는 그룹으로 연령군을 정의할 수 있는데, 예를 들어 그룹 1은 0 내지 5세로 정의하고, 그룹 2는 6 내지 15세로 정의할 수 있다. 그룹 1 및 그룹 2는 연령층이 다르기 때문에 시물레이션 단위 일자당 연령의 증가율이 다르다.

[0029] 여기서, 백신접종에 관한 정보는 시물레이션 기간, 백신접종이 시행되는 시기, 백신효능, 접종율 및 백신을 시행할 그룹 등을 나타낸다. 따라서, 연령-특이 산술모듈(123)은 그룹 1 또는 그룹 2 등의 연령군에서의 감염된 사람수를 출력할 수 있다. 예를 들어, 시물레이션 기간으로 3650가 입력되고 백신접종이 시행되는 시기로 365가 입력된다면 총 10년의 시물레이션 기간에서 1년이 되는 시점에 백신접종이 시작됨을 의미한다. 예를 들어, 백신효능이 80%라는 것은 백신을 접종한 사람이 접종하지 않는 사람에 비해 감염이 될 위험이 80% 감소한다는 것을 의미한다. 접종률이 50% 입력된다면 전체 인구중 50%가 백신을 접종하였다는 것을 의미한다. 백신을 시행할 그룹은 연령-특이 산술모듈(123)에서는 연령-특이적이기 때문에 연령별로 그룹을 지정할 수 있으며, 백신을 어느 그룹에 접종하게 되는지를 입력하는 변수이다.

[0030] 비례 산술모듈(125)은 연령-특이 산술모듈(123)의 입력 변수에 출생률을 추가하여 입력받고 다양한 변수를 포함하는 미분방정식을 이용하여 감수성(susceptible)집단의 연령군에서의 질병 발생수를 출력한다. 따라서, 비례 산술모듈(125)에서는 질병에 관한 대처효과가 출생률을 통해 반영된다.

[0031] 통계 발생예측부(130)의 ARIMA 모듈(131), 개입 모듈(133), 전이 함수 모듈(135), 개입 전이 함수 모듈(137) 및 인플루엔자 모듈(139)에서는 입력값이 차이가 있으며, 시간에 따른 전염병 발생수 및 그 예측값을 출력한다.

[0032] 다음 표 1은 통계 발생예측부(130)의 각 모듈에 따른 입력 변수를 보여준다.

표 1

	시간	전염병 발생수	전이 함수 변수	개입 변수
ARIMA 모듈	○	○		
개입 모듈	○	○		○
전이 함수 모듈	○	○	○	
개입 전이 함수 모듈	○	○	○	○
인플루엔자 모듈	○	○		

[0034] 도 2는 통계 발생예측부의 입력 변수의 일례를 보여주는 화면이다.

[0035] 도 2를 참조하면, 첫번째 열은 시간을 나타내고 두번째 열은 그 해(year)에 발생한 백일해 발생수를 나타내고

세번째 열은 출생률을 나타내며 네번째 열은 사망률을 나타낸다. 세번째 및 네번째 열은 전이 함수에 이용되는 변수들이다.

[0036] ARIMA 모듈(131)은 현재 관측값과 과거의 관측값 간에 상관성이 존재하는 시계열 자료를 분석하며 시계열 자료에 추세나 이분산성이 존재하면 로그변환이나 차분을 이용하여 안정화시킬 수 있고, 시간 및 전염병 발생수를 입력받아 다음 수학적 식 1을 이용하여 시간에 따른 전염병 발생수 및 그 예측값을 출력한다.

수학적 식 1

$$(1 - B)^d z_t = \frac{1 - \theta_1 B - \dots - \theta_p B^p}{1 - \phi_1 B - \dots - \phi_p B^p} a_t = \frac{\theta(B)}{\phi(B)} a_t$$

[0037]

[0038] 여기서, p 및 q는 자동 회귀(auto regressive: AR)와 이동 평균(moving average: MA)의 차수이며 d는 차분의 차수이다. $\phi_1, \phi_2 \dots \phi_p$ 및 $\theta_1, \theta_2 \dots \theta_q$ 는 자동 회귀 및 이동 평균의 상관계수를 나타내고, a_t 는 백색과정을 의미하며, B는 $B^k z_t = z_{t-k}$ 를 만족하는 후진연산자이다. 자료가 계절성을 포함하고 있지 않다면 계절성을 나타내는 항은 무시할 수 있다.

[0039] 개입(intervention) 모듈은 개입이 발생했을 경우 그 개입변수가 결과에 어떤 영향을 미치는지 관찰하는 것으로 개입변수는 다음 수학적 식 2인 펄스 함수와 다음 수학적 식 3인 계단 함수로 이루어지고, 시간, 전염병 발생수 및 개입 변수를 입력받아 시간에 따른 전염병 발생수 및 그 예측값을 출력한다.

수학적 식 2

$$P_t^{(T)} = \begin{cases} 0, t = T \\ 1, t \neq T \end{cases}$$

[0040]

[0041] 수학적 식 2는 어느 특정한 시점에서만 영향을 주는 함수이다.

수학적 식 3

$$S_t^{(T)} = \begin{cases} 0, t < T \\ 1, t \geq T \end{cases}$$

[0042]

[0043] 수학적 식 3은 T시간 이후에도 개입이 계속 영향을 주는 함수이다.

[0044] ARIMA모듈에서 분석결과에 영향을 주는 외부 변수가 존재한다면 이를 분석에 반영하는 것이 필요하다. 이러한 경우 분석하고자 하는 시계열을 출력계열(output series)이라 하고 출력계열을 설명하고자 하는 변수를 입력계열(input series)이라 한다. ARIMA모듈에서 외부 변수를 더 고려한 것이 전이 함수 모듈(135)이며, 시간, 전염병 발생수 및 외부 변수인 전이 함수 변수를 입력받아 시간에 따른 전염병 발생수 및 그 예측값을 출력한다. 입력계열과 출력계열은 다음 수학적 식 4로 연관되어진다.

수학식 4

$$y_t = \gamma(B)x_t + n_t$$

[0045]

[0046]

여기서, x_t 는 입력계열, y_t 는 출력계열, $\gamma(B)$ 는 전이함수 및 n_t 는 잡음계열을 나타낸다.

[0047]

개입 전이 함수 모듈(137)은 시간, 전염병 발생수, 개입 변수 및 외부 변수인 전이 함수 변수를 입력받아 개입 모듈(133) 및 전이 함수 모듈(135)의 상기 수식을 이용하여 시간에 따른 전염병 발생수 및 그 예측값을 출력한다.

[0048]

인플루엔자 모듈(139)은 시간 및 전염병 발생수를 입력받아 회귀모형으로 시간, 시간에 대한 사인 및 코사인 함수, 온도, 휴일효과, 겨울방학효과 등을 독립변수로 이용하여 시간에 따른 전염병 발생수 및 그 예측값을 출력한다.

[0049]

출력부(140)는 LCD(liquid crystal display), LED(Light Emitting Diode) 디스플레이, PDP(Plasma Display Pannel) 및 OLED(Organic Light Emitting Diodes) 디스플레이 중 적어도 하나일 수 있다.

[0050]

본 실시예에 따른 질병 발생 예측 시스템은 미분방정식을 구현할 수 있는 모델메이커(ModelMaker) 및 SAS 프로그램의 기능을 R 언어로 구현한 후, C 계열의 언어를 이용하여 인터페이스를 구성할 수 있다. 예를 들어, C 계열의 언어 예를 들어, C 언어, C++, C#(sharp) 중 C#를 이용하여 본 실시예에 따른 질병 발생 예측 시스템의 인터페이스를 구성할 수 있다. R 언어 및 C 계열의 언어의 연동을 위해서는 COM(Component Object Model)을 이용하여, C 계열의 언어에서 개발된 윈도우 프로그래밍을 지원하기 위하여 COM 인터페이스를 사용함으로써 R언어에서 사용가능하다.

[0051]

도 3은 본 실시예에 따른 질병 발생 예측 시스템의 개발 과정을 설명하는 도면이다.

[0052]

모델메이커(ModelMaker)는 ModelKinetix에서 개발한 것으로, 모델링(modeling)을 하는 프로그램이다. 모델메이커는 모델링을 필요로 하는 환경과학, 화학, 생리학, 사회학, 전염병학, 약물동력학, 생태학, 수학 등 다양한 학문분야에 적용할 수 있다. 모델메이커의 광범위한 함수들을 이용하여 여러 시스템에 대한 모델링을 할 수 있으며, 이 결과들을 최적화(optimization), 최소화(minimization), 몬테칼로 시뮬레이션(Monte Carlo simulation), 민감도 분석(sensitivity analysis)을 이용하여 합리적으로 설명할 수 있다.

[0053]

SAS(Statistical Analysis system 또는 Strategic Applications Software)는 SAS 연구소에서 개발한, 그리고 개발이 진행 중인 각종 통계 분석을 위한 통계 패키지 프로그램이다. 이러한 SAS는 여러 소프트웨어 및 기타 지원 소프트웨어로 구성되어있다.

[0054]

R은 벨 연구소에서 개발된 S언어의 확장으로 공개용 소프트웨어이다. R은 학계에서는 하나의 표준 프로그램 언어로 사용되고 있는 실정이다. R언어를 이용하여 데이터 분석을 위한 프로그램을 개발하고 이를 패키지화하여 R에 포함시키고 있어 기존의 통계패키지에서는 사용할 수 없는 최신의 통계기법을 R에서 사용할 수 있다는 장점이 있다.

[0055]

C# 언어는 새로운 Microsoft .NET 플랫폼에서 동작되는 애플리케이션을 빠르고, 쉽게 구축할 수 있게 해주는 객체지향 언어이다. 이러한 C#의 장점은 윈도우 API를 사용할 수 있다는 것이다. 운영체제는 수많은 API(Application Programming Interface)로 이루어져 있는데, API란 운영체제를 구성하고 있는 수많은 부분을 의미하며, 그 자신 또한 프로그램이다. 이러한 API를 프로그래밍에 활용하면 해당 운영체제에서의 애플리케이션을 쉽게 만들 수 있다.

[0056]

도 4는 본 실시예에 따른 질병 발생 예측 시스템의 초기 입력인터페이스 화면이다. 도 4를 참조하면, 이용하고자 하는 모듈을 클릭하여 선택한 후 예측에 필요한 변수를 설정하여 질병 발생을 예측할 수 있다.

[0057]

이하에서는, 산술 발생예측부(120)의 여러 모듈 중 연령-특이 산술모듈(123)을 예로 들어 실행을 설명하기로 한다.

[0058]

도 5는 본 실시예에 따른 연령-특이 산술모듈의 변수 설정 화면이고, 도 6은 출력화면이다.

[0059]

도 5 및 도 6을 참조하면, 전염병학에 관한 정보, 예를 들어, 그룹에 따른 연령 증가율(aging rate)와 잠재기

및 백신접종에 관한 정보를 입력할 수 있고, 도 6의 오른쪽 상단 플롯에 연령 그룹에 대한 플롯이 출력되고 하단에는 결과가 출력된다. 하단의 결과창 옆에 'Plot R value'(610) 버튼을 클릭하면 R값에 대한 플롯만을 그려 주고, 'Export'(630)를 클릭하면 결과를 파일 형태가 CSV인 파일로 저장한다.

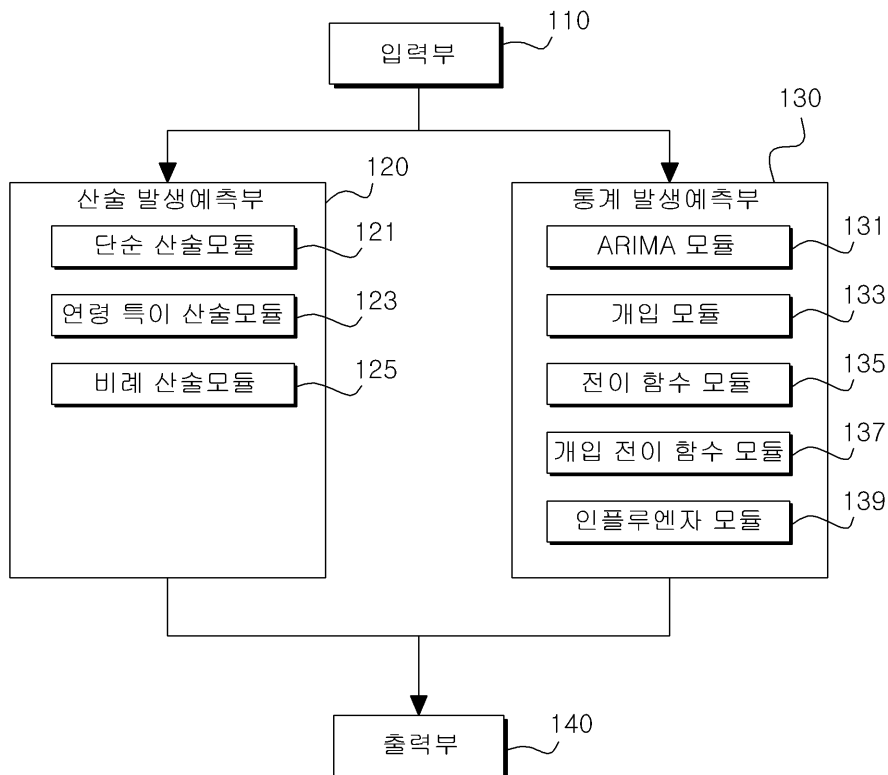
- [0060] 이하에서는, 통계 발생예측부(130)의 여러 모듈 중 개입 전이 함수 모듈(137)을 예로 들어 실행을 설명하기로 한다.
- [0061] 도 7은 본 실시예에 따른 개입 전이 함수 모듈의 변수 설정 화면이고, 도 8은 출력화면이다.
- [0062] 도 7 및 도 8을 참조하면, 펄스 함수, 계단 함수 및 전이 변수의 세팅이 나타나며 전염병 발생수를 입력할 수 있고 로그 변환 및 차분 차수를 입력할 수 있으며, 도 8에서는 시간에 따른 질병 발생수의 그래프가 출력된다.
- [0063] 본 실시예에 따른 질병 발생 예측 시스템의 출력 화면은 표 또는 그래프의 플롯으로 표시된다. 표에는 회귀계수, 검정통계량 및 설명력 등이 포함되며, 플롯의 줌 기능을 포함하여 마우스를 이용하여 자세히 보고 싶은 부분을 드래그하면 해당 부분이 확대되어 출력된다. 출력된 플롯에서 분석값들을 볼 수 있다.
- [0064] 본 실시예에서 사용되는 '~부'라는 용어는 소프트웨어 또는 FPGA(field-programmable gate array) 또는 ASIC과 같은 하드웨어 구성요소를 의미하며, '~부'는 어떤 역할들을 수행한다. 그렇지만 '~부'는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '~부'는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '~부'는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들, 및 변수들을 포함한다. 구성요소들과 '~부'들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 '~부'들로 결합되거나 추가적인 구성요소들과 '~부'들로 더 분리될 수 있다. 뿐만 아니라, 구성요소들 및 '~부'들은 디바이스 또는 보안 멀티미디어카드 내의 하나 또는 그 이상의 CPU들을 재생시키도록 구현될 수도 있다.
- [0065] 상술한 실시예에서 사용한 구체적인 수치 또는 도면에 나타난 수치는 본 발명의 일 실시예를 설명하기 위한 것에 불과하므로 본 발명의 내용이 이러한 구체적인 수치에 한정되지 않는다.
- [0066] 상술한 모든 기능은 상기 기능을 수행하도록 코딩된 소프트웨어나 프로그램 코드 등에 따른 마이크로프로세서, 제어기, 마이크로제어기, ASIC(Application Specific Integrated Circuit) 등과 같은 프로세서에 의해 수행될 수 있다. 상기 코드의 설계, 개발 및 구현은 본 발명의 설명에 기초하여 당업자에게 자명하다고 할 것이다.
- [0067] 이상 본 발명에 대하여 실시예를 참조하여 설명하였지만, 해당 기술 분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시켜 실시할 수 있음을 이해할 수 있을 것이다. 따라서 상술한 실시예에 한정되지 않고, 본 발명은 이하의 특허청구범위의 범위 내의 모든 실시예들을 포함한다고 할 것이다.

부호의 설명

- | | | |
|--------|---------------|------------------|
| [0068] | 110: 입력부 | 120: 산술 발생예측부 |
| | 121: 단순 산술모듈 | 123: 연령-특이 산술모듈 |
| | 125: 비례 산술모듈 | 130: 통계 발생예측부 |
| | 131: ARIMA 모듈 | 133: 개입 모듈 |
| | 135: 전이 함수 모듈 | 137: 개입 전이 함수 모듈 |
| | 139: 인플루엔자 모듈 | 140: 출력부 |

도면

도면1



도면2

도면2는 'Import Wizard ...' 창의 'Data Verification' 탭을 보여줍니다. KCDC 로고가 오른쪽 상단에 있습니다. 중앙에는 데이터 테이블이 표시되어 있으며, 'year' 열이 선택되어 있습니다. 오른쪽에는 'Data count'와 'Data columns' 섹션이 있습니다.

	year	pertussis	birth	death
▶	1970	3818	31.2	8
	1971	1436	31.2	7.2
	1972	989	28.5	6.3
	1973	1846	28.4	7.9
	1974	3077	26.6	7.2
	1975	1182	24.8	7.7
	1976	1177	22.2	7.5
	1977	3162	22.7	6.9
	1978	713	20.4	6.8
	1979	787	23	6.4

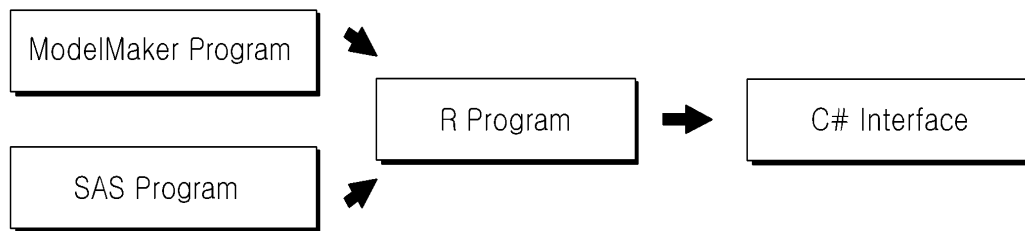
Data count: Total index size : 36

Data columns:

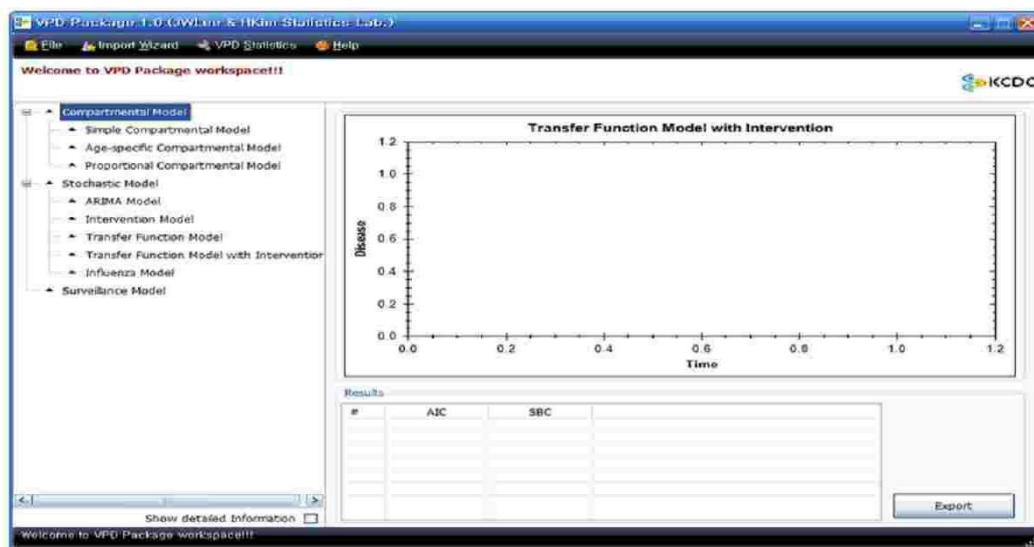
- ☒ year
- ☒ pertussis
- ☒ birth
- ☒ death

Bottom buttons: < Back, Finish, Cancel

도면3



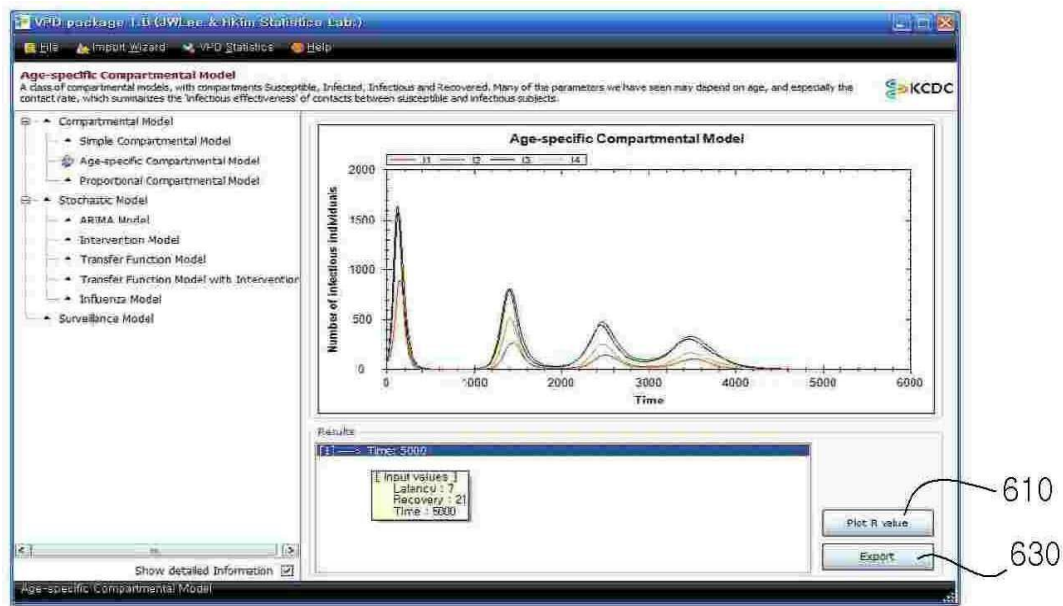
도면4



도면5

The screenshot shows the 'Configuration for VPD package' dialog box for the 'Age-specific Compartmental Model'. It includes sections for 'Information of Epidemiology', 'Information of Vaccination', and 'Initial value for compartments', each with a text input field. The 'OK' and 'Cancel' buttons are at the bottom.

도면6



도면7

The screenshot shows the 'Configuration for VPD package' dialog box. The title bar reads 'Configuration for VPD package'. The main section is titled 'Transfer Function Model with Intervention Configuration'. It contains two main sections: 'Variable setting' and 'Transformation & Differentiation'. In the 'Variable setting' section, 'Disease' is set to 2, 'Step function setting' is set to ..., 'Pluse function setting' is set to ..., and 'Transfer variable setting' is set to In the 'Transformation & Differentiation' section, 'Log transformation' is selected, 'Power transformation' is set to 1, 'Differentiation' is set to 1, and 'Forecastion' is set to 2. At the bottom, there are 'OK' and 'Cancel' buttons.

도면8

