ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Введение

Процесс изменения температуры характеризуется множеством скрытых параметров, таких как конвекционные потоки, солнечная активность, деятельность человека и другие, большинство из которых сложно прогнозируемы или вовсе неизвестны, поэтому в данной работе рассматривается динамика изменения температуры без попыток установления причин ее изменения. Такой подход применяется в настоящее время и позволяет получить более высокую точность в долгосрочной перспективе с потерей точности в краткосрочной. Однако наиболее используемым методом анализа является усреднение величины по времени, то есть, прогнозируемая среднемесячная температура в июле следующего года рассчитывается как среднее арифметическое среднемесячной температуры в июле за последние несколько десятилетий. Данная работа предлагает альтернативный метод: анализ поэтапного изменения среднемесячной температуры с шагом в один месяц за последние двести лет, получение функциональной зависимости среднемесячной температуры от времени, и прогноз последующих значений с помощью временных рядов.

Для прогноза были использованы три модели: первая модель представляет собой гармоническую функцию $f(t)$; вторая модель – ту же самую функцию, но с добавлением случайного слагаемого, ответственного за флуктуацию значений измеряемой температуры – с помощью него мы вносим поправки на различие теоретической функции и реального процесса; третья модель представляет собой дискретную функцию с шагом в один месяц, значения которой являются экспериментально измеренными значениями (дискретность функции не влияет на алгоритм прогнозирования, поскольку временной ряд априори является дискретным).

Подбор теоретической функции для прогнозирования

Для получения функциональной зависимости среднемесячной температуры от месяца (и года) измерения использовались данные с сайта летописи погоды в Москве [6]. Диапазон данных принят с января 1823 года по декабрь 2022 года. С помощью приложения Curve Fitting Toolbox программы MATLAB [Приложение 1] была подобрана функция, описывающая данный процесс с достаточно хорошей точностью ($R^2 = 0.9352, RMSE = 2.653$). Полученная функция имеет вид:

$$f(t) = a_0 t + a_1 sin(b_1 t + c_1) + a_2 sin(b_2 t + c_2) + a_3 sin(b_3 t + c_3)$$

Таблица 4. Значения коэффициентов функции $f(t)$

a_0	a_1	b_1	c_1	a_2	b_2	c_2	a_3	b_3	c_3
0.004	14.16	0.5236	−2.152	2.622	0.001531	1.625	0.1462	0.02867	−31.08

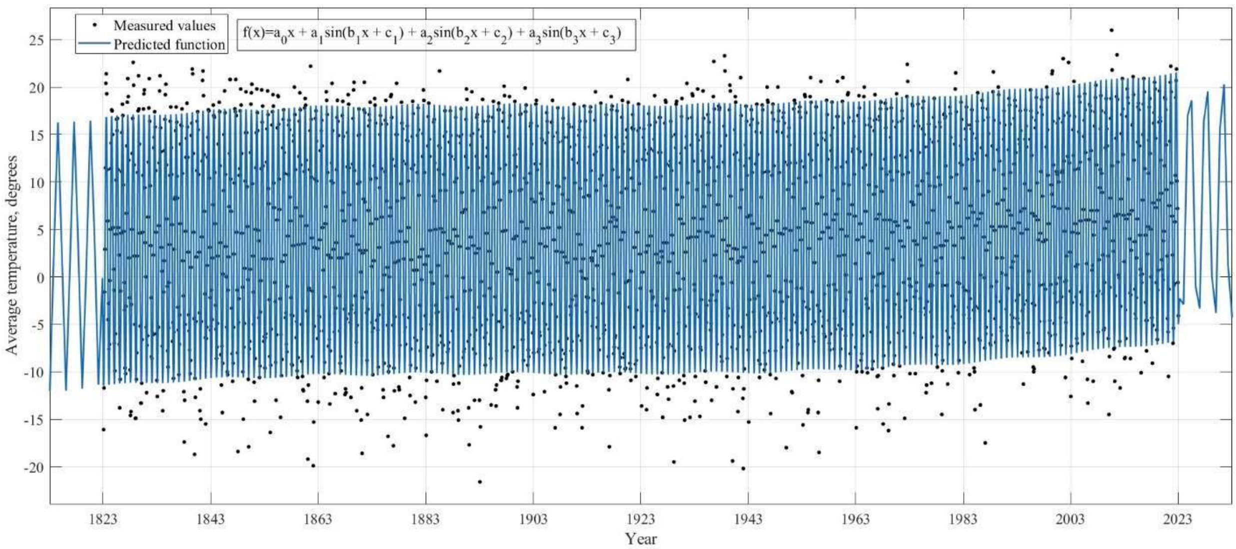

Рис. 3. Измеренные значения и график функции $f(t)$

Следует обратить внимание на первое слагаемое $a_0 t$, являющееся линией тренда данной функции. Положительный коэффициент $a_0 = 0.004$ свидетельствует о линейном росте средней температуры с течением времени или, так называемом, глобальном потеплении.

Прогнозирование функции $f(t)$

С помощью программы MATLAB [Приложение 2] было произведен прогноз заданной функции $f(t)$ методом построения временного ряда длиной в 2400 значений по 600 заданным значениям с шагом $\Delta t = 1$.

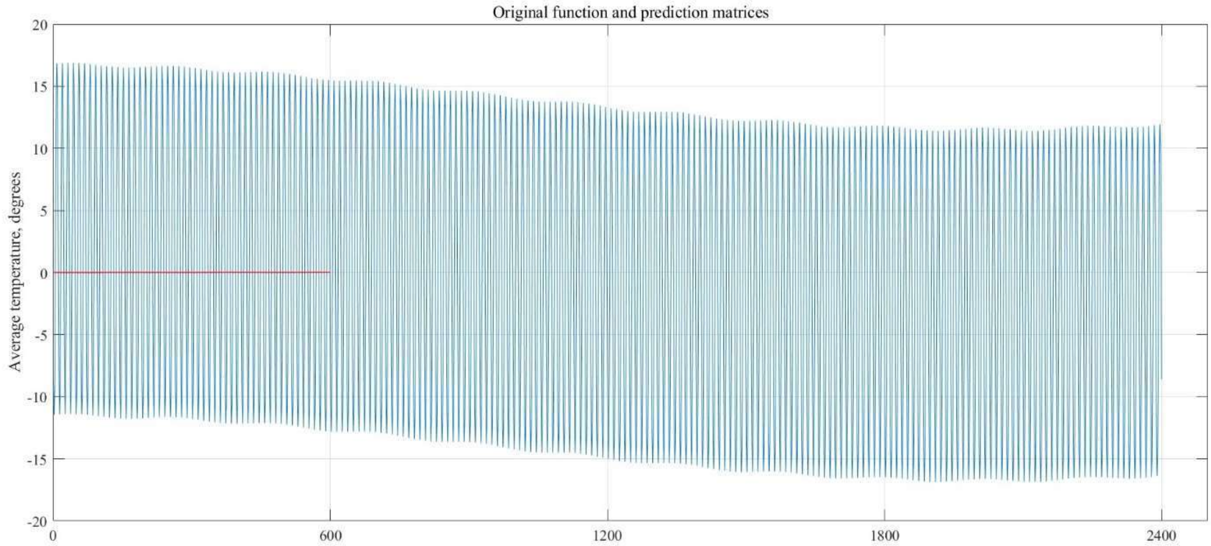

Рис. 4. График исходной функции и матрицы прогноза (t = 2400, Δt = 1, w = 600)

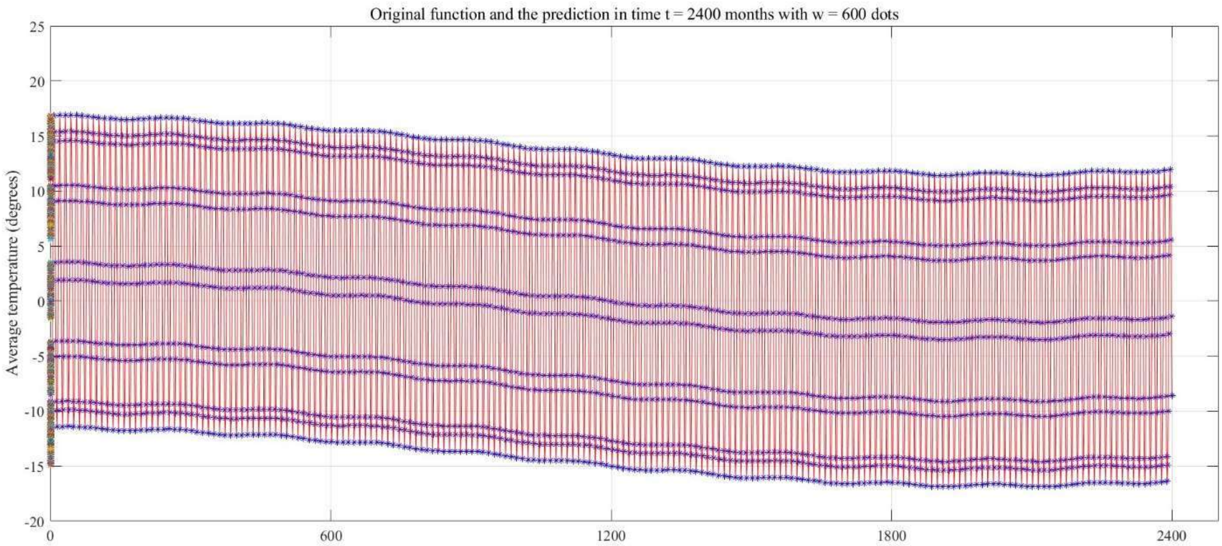

Рис. 5. График исходной функции (синие звёздочки), совмещенный с графиком прогноза (тонкая красная линия) (t = 2400, Δt = 1, w = 600)

Максимальная погрешность при данном прогнозировании имеет порядок 10^{-8}, что говорит о высокой точности прогноза.

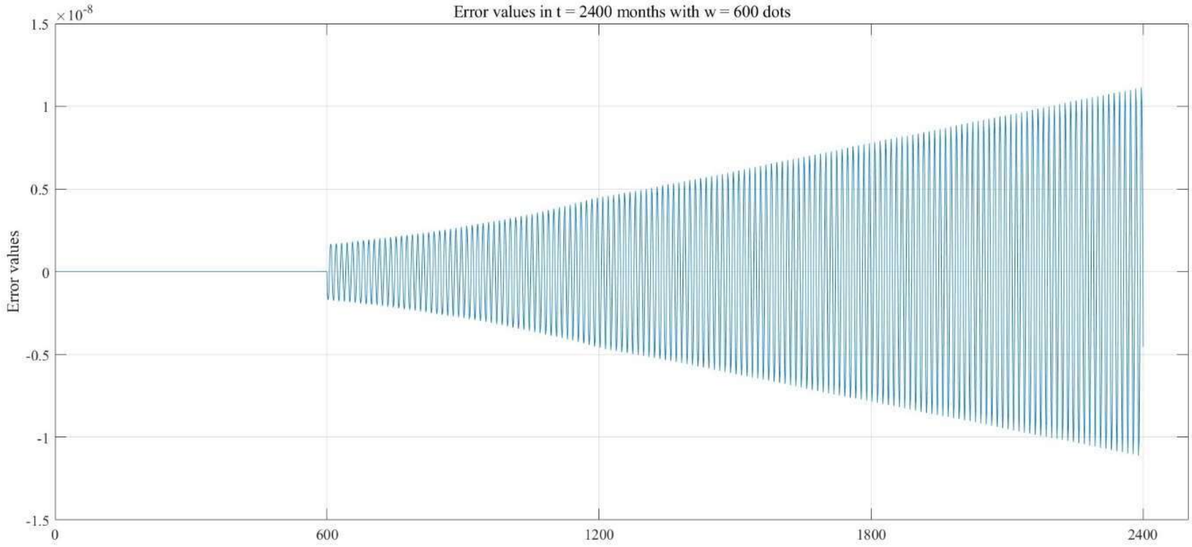

Рис. 6. График погрешности прогноза функции (t = 2400, Δt = 1, w = 600)

Для лучшей визуализации можно представить график следующим образом: по оси абсцисс отложить временной промежуток с 1973 по 2173 года. Период с 1973 по 2022 год является данными для построения прогноза, а период с 2023 по 2173 – прогнозируемым. Тогда максимальная погрешность в предсказании температуры к 2173 году составит лишь 10^{-8} градуса.

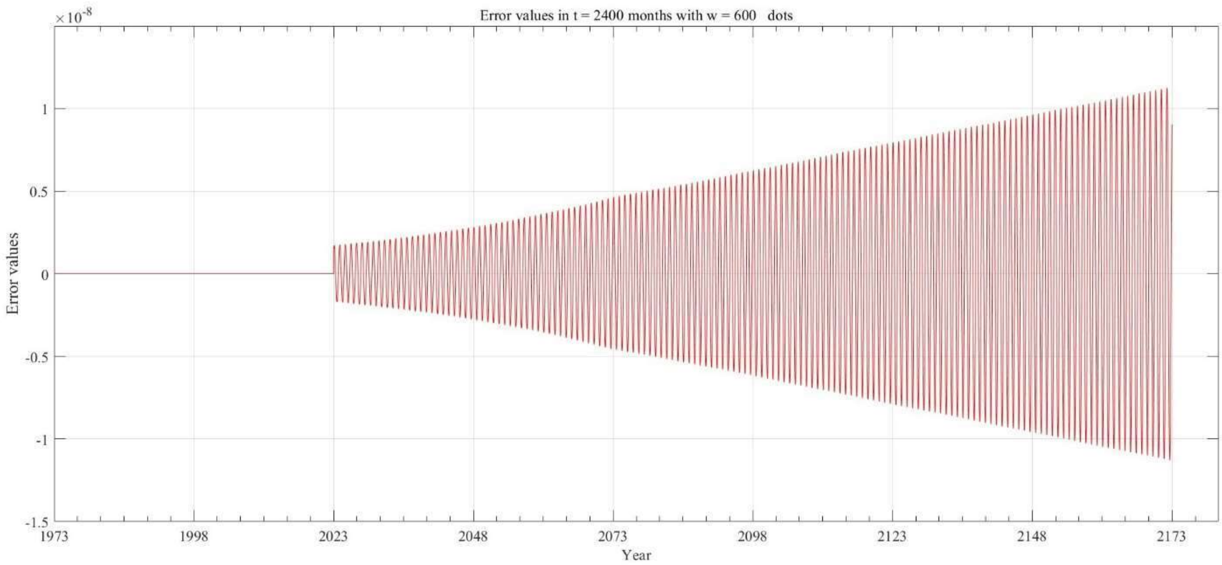

Рис. 7. График погрешности прогноза функции по годам (t = 2400, Δt = 1, w = 600)

Прогнозирование функции $f_{err}(t)$

Вторая модель представляет собой сумму функции $f(t)$ и случайного слагаемого $k\ rand(1)$:

$$f_{err}(t) = a_0 t + a_1 sin(b_1 t + c_1) + a_2 sin(b_2 t + c_2) + a_3 sin(b_3 t + c_3) + k\ rand(1)$$

Коэффициент k является амплитудой добавочного случайного значения, и в данном контексте может интерпретироваться как цена деления измерительного прибора. Для прогноза было выбрано значение $k = 0.05$.

Максимальная погрешность при данном прогнозировании составляет $0.119°$, что на семь порядков больше, чем для функции без случайного слагаемого.

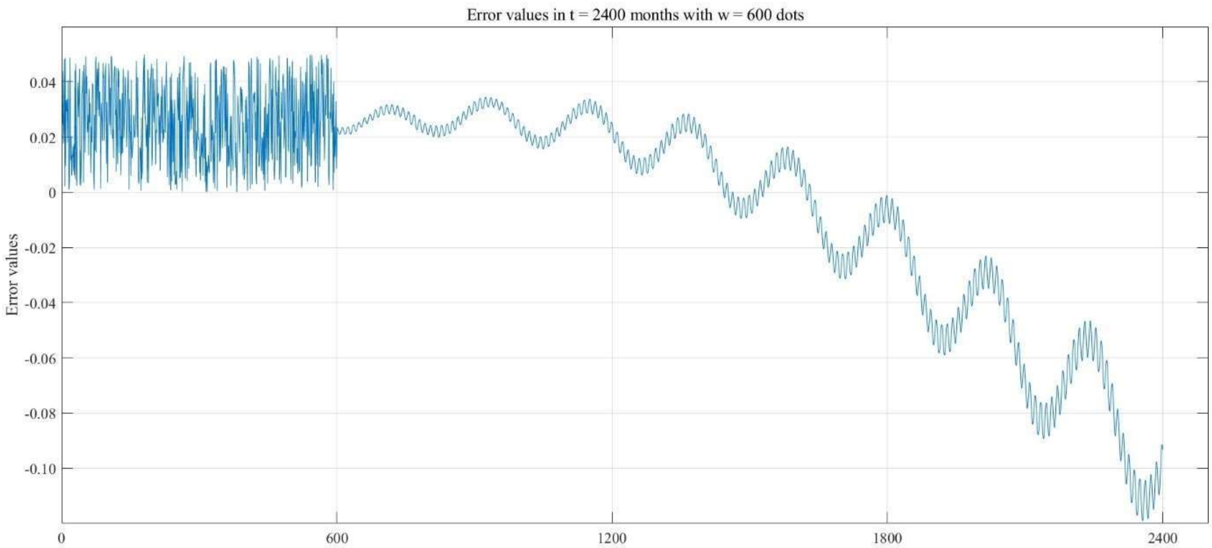

Рис. 8. График погрешности прогноза функции с добавочным случайным слагаемым (t = 2400, Δt = 1, w = 600, k = 0.05)

Прогнозирование функции $f_{exp}(t)$

Третья модель представляет собой дискретную функцию с шагом в один месяц, значения которой являются экспериментально измеренными значениями, взятыми с сайта летописи погоды в Москве [6].

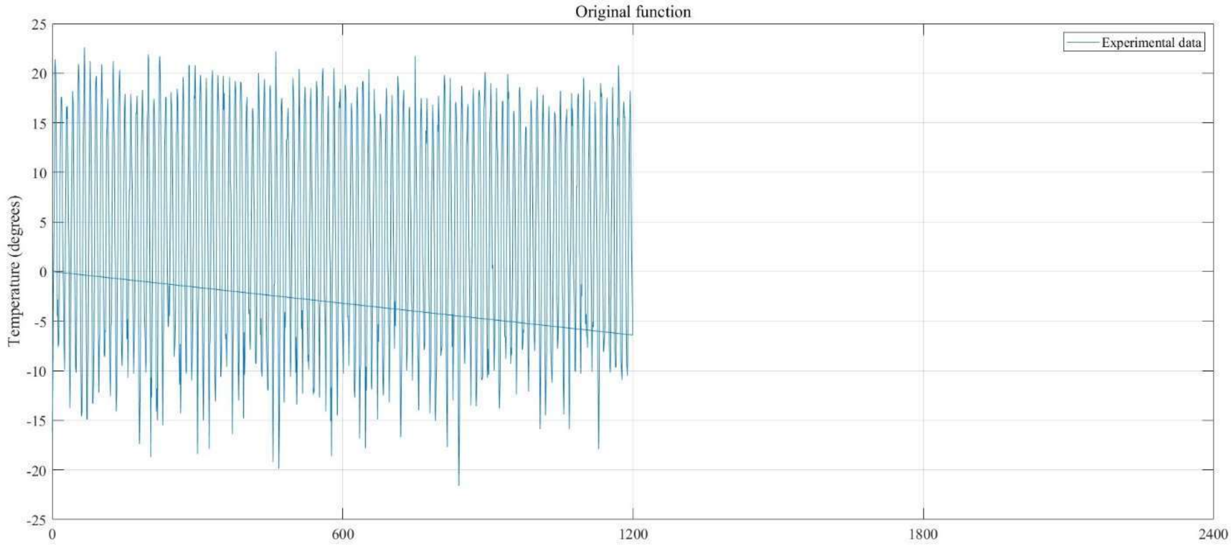

Рис. 9. График исходной экспериментальной функции (t = 2400, Δt = 1, w = 600)

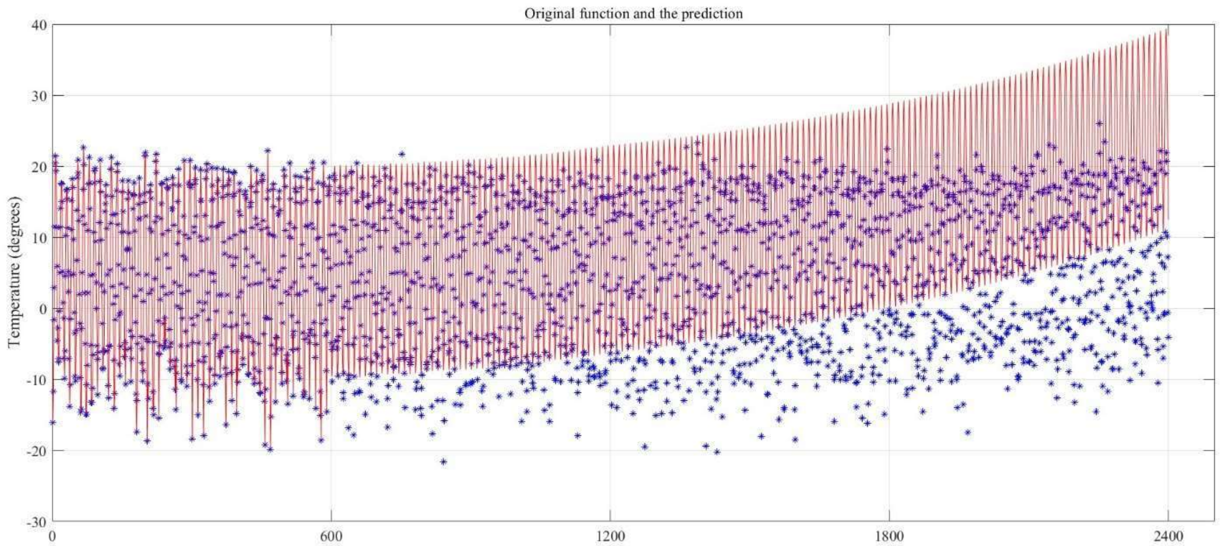

Рис. 10. График исходной экспериментальной функции (синие звёздочки), совмещенный с графиком прогноза (тонкая красная линия) (t = 2400, Δt = 1, w = 600)

На графике видно, что предсказанные значения температуры достаточно быстро начинают отличаться от реальных.

Максимальная погрешность при данном прогнозировании составляет 23.73°, что на девять порядков больше, чем для функции $f(t)$.

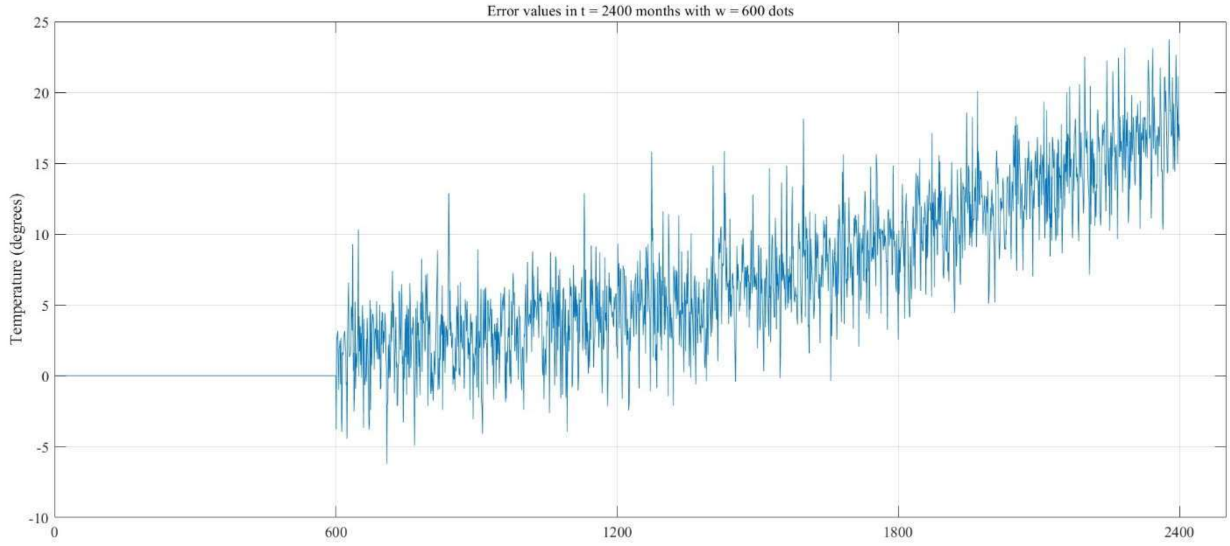

Рис. 11. График погрешности прогноза экспериментальной функции

(t = 2400, Δt = 1, w = 600)

Оценка погрешностей прогноза для различных *w, t* и Δt

Таблица 5. Погрешности при изменении w и t $(\Delta t = 1)$

Количество исходных точек	w = 600	w = 700	w = 800	w = 900	w = 1000
Погрешность прогноза функции Δf					
t = 2000 месяцев	$8.9122 * 10^{-9}$	$8.1030 * 10^{-9}$	$1.9464 * 10^{-7}$	$6.4872 * 10^{-9}$	$5.4429 * 10^{-9}$
t = 2200 месяцев	$1.0034 * 10^{-8}$	$9.1927 * 10^{-9}$	$2.2154 * 10^{-7}$	$7.3523 * 10^{-9}$	$6.1733 * 10^{-9}$
t = 2400 месяцев	$1.1139 * 10^{-8}$	$1.0229 * 10^{-8}$	$2.4876 * 10^{-7}$	$8.3019 * 10^{-9}$	$6.9537 * 10^{-9}$
Погрешность прогноза функции с помехами Δf_{err} $(k = 0.05)$					
t = 2000 месяцев	0.0500	0.0500	0.0500	0.0500	0.0500
t = 2200 месяцев	0.0781	0.0500	0.0500	0.0500	0.0537

t = 2400 месяцев	0.1190	0.0500	0.0500	0.0500	0.0542
Погрешность прогноза экспериментальных значений Δf_{exp}					
t = 2000 месяцев	20.1039	$2.4748 * 10^6$	$5.6039 * 10^4$	13.8849	$3.0095 * 10^4$
t = 2200 месяцев	22.5121	$2.3240 * 10^7$	$3.1920 * 10^5$	14.5788	$2.5215 * 10^5$
t = 2400 месяцев	23.7300	$2.3374 * 10^8$	$1.8449 * 10^6$	15.0341	$1.9929 * 10^6$

Таблица 6. Погрешности при изменении w и Δt $(t = 2400)$

Количество исходных точек	w = 600	w = 700	w = 800	w = 900	w = 1000
Погрешность прогноза функции Δf					
Δt = 0.1 месяца	0.0617	$5.0555 * 10^{-11}$	$3.1085 * 10^{-4}$	$7.4884 * 10^{-5}$	$4.6233 * 10^{-5}$
Δt = 0.5 месяца	$6.0019 * 10^{-9}$	$9.9629 * 10^{-8}$	$2.4771 * 10^{-8}$	$1.2676 * 10^{-8}$	$8.7284 * 10^{-8}$
Δt = 1 месяц	$1.1139 * 10^{-8}$	$1.0229 * 10^{-8}$	$2.4876 * 10^{-7}$	$8.3019 * 10^{-9}$	$6.9537 * 10^{-9}$
Δt = 2 месяц	$4.4473 * 10^{-9}$	$6.3540 * 10^{-8}$	$5.2108 * 10^{-9}$	$1.1747 * 10^{-9}$	$6.2681 * 10^{-8}$

Вывод

В общем случае точность прогноза увеличивается с увеличением количества исходных точек и уменьшается с увеличением прогнозируемого периода. Однако в случае экспериментальной функции существуют значения $w = 300 * n,\ n \in \mathbb{N}$, кратные периодам исходных функций, при которых погрешность намного меньше, чем при остальных значениях w.

В случае функции с помехами при достаточно большом количестве исходных данных ($w \geq 700$) максимальная погрешность становится постоянной и равной амплитуде добавочного случайного слагаемого.

Зависимость точности прогноза от шага Δt имеет более сложную природу.

Использование экспериментальной функции для прогноза на практике не применимо, поскольку погрешность при использовании данного метода составляет порядок прогнозируемой величины, что свидетельствует о наличии скрытых параметров, которые построенная модель не может учесть.

Наиболее подходящей моделью является прогнозирование функции с помехами: в отличие от экспериментальной функции, погрешность при ее прогнозировании имеет относительно небольшое значение, а внесение случайного слагаемого позволяет приблизить данную модель к реальному процессу с большей точностью, нежели модель без помех.

В этом случае итоговая погрешность складывается из погрешности при моделировании функции и погрешности при ее прогнозировании.