



Московский государственный университет имени М.В.Ломоносова
Факультет вычислительной математики и кибернетики
Кафедра математической статистики

Лазар Владислав Игоревич

**Анализ и применение многофакторных моделей
динамики лекарственных веществ в медицинских исследованиях**

Выпускная квалификационная работа

Научный руководитель:
доцент, кандидат физико-математических наук
Захарова Татьяна Валерьевна

Москва, 2025

Содержание

1	Введение	3
2	Постановка задачи	5
3	Описание данных	7
4	Описание моделей	9
4.1	PBFTPК	9
4.2	Кусочная модификация PBFTPК	10
4.3	EPBFTPК	11
4.4	Функции потерь	11
4.5	Описание итогового алгоритма	12
5	Полученные результаты	13
6	Заключение	20
	Список литературы	21

1 Введение

Фармакокинетика (ФК) — это наука, которая описывает всасывание, распределение, метаболизм и выведение (ADME) лекарственных средств из организма. Её одной из ключевых задач является количественная оценка того, как действующее вещество лекарственного препарата перемещается в различных тканях и органах под влиянием множества факторов: физиологических, биохимических, генетических и внешних (например, сопутствующее лечение, образ жизни пациента и т. д.).

В последние годы всё большую популярность приобретают многофакторные модели, учитывающие различные аспекты сложного взаимодействия между лекарством и организмом, а также позволяющие более точно описывать межиндивидуальную вариабельность. Одним из перспективных направлений в этой области считаются модели типа **PВFТРК** [1], которые обладают богатым набором параметров и могут учитывать индивидуальные особенности организма, в том числе пол, возраст, массу тела, функциональное состояние органов и т. д.

Однако при исследовании фармакокинетических свойств лекарственных средств возникает ряд проблем, связанных с анализом данных:

1. **Малые выборки.** В фармакологических исследованиях нередко сталкиваются с ограничением по числу доступных испытуемых или экспериментальных данных. Клинические испытания часто подразумевают небольшие группы пациентов, что затрудняет статистически достоверную оценку модели.
2. **Неоднородность данных по времени.** Измерения концентрации лекарственных веществ в крови или тканях проводятся с разными интервалами, и при этом часть данных может отсутствовать (например, пропущенные точки). Это приводит к неравномерным временным рядам и усложняет анализ.
3. **Сильная зашумлённость данных.** Биологические процессы обладают существенной естественной вариабельностью, а методики измерения могут вносить дополнительные погрешности. В результате концентрационные кривые оказываются сильно зашумлёнными.
4. **Большое количество выбросов.** В фармакологических данных часто встречаются аномальные значения, которые выходят за рамки ожидаемого диапазона.

Такие выбросы могут искажать результаты анализа и повышают риски некорректных выводов.

Для корректной интерпретации результатов, получаемых с помощью многофакторных моделей, важно учитывать все перечисленные особенности. В данной работе будут рассмотрены концептуальные основы моделирования фармакокинетических процессов с акцентом на применение РВФТРК-подхода. Особое внимание будет уделено статистическому анализу и методам предобработки данных, направленным на снижение влияния шума и выбросов, а также корректную работу с малым объёмом данных и их временной неоднородностью.

Дальнейшие главы работы будут описывать современные методики построения многофакторных моделей, принципы адаптации РВФТРК к конкретным клиническим сценариям, а также практически применимые техники валидации полученных результатов. Ожидается, что рассмотренные подходы позволят более детально и адекватно описывать фармакокинетические процессы, учитывая всю сложность взаимодействия биологических и внешних факторов.

2 Постановка задачи

Классическая РВФТРК-модель предполагает, что концентрация лекарственного вещества в организме возрастает до некоторого момента, а затем монотонно убывает. На практике, однако, нередко наблюдаются более сложные траектории, которые могут включать повторные пики, периоды плато, перемены скорости элиминации и другие особенности кривой концентрации. При этом сохраняется важное требование — физическая интерпретируемость модели, позволяющая связать параметры с реальными физиологическими процессами (кровоток, метаболические пути, связывание с белками и др.).

С учётом перечисленных в **Введении** проблем (малые выборки, неоднородность данных по времени, наличие шума и выбросов), **цель данной работы** заключается в построении и исследовании **модифицированной РВФТРК-модели**, которая:

- Способна описывать неоднократные изменения концентрации (в том числе повторные подъёмы и спады).
- Устойчиво обрабатывает зашумлённые и неравномерные данные с выбросами, сохраняя корректность оценок параметров.
- Остаётся физически интерпретируемой, т. е. параметры и структура модели отражают реальные процессы распределения и метаболизма лекарственного вещества.
- Может применяться при малом объёме экспериментальных данных без существенной потери в точности.

Таким образом, основная **задача** работы — разработать и обосновать набор математических и вычислительных приёмов, позволяющих перейти от классического упрощённого описания фармакокинетического процесса к более общей и гибкой модели РВФТРК, учитывающей реальные особенности фармакодинамики лекарственных веществ. В рамках этой задачи предполагается:

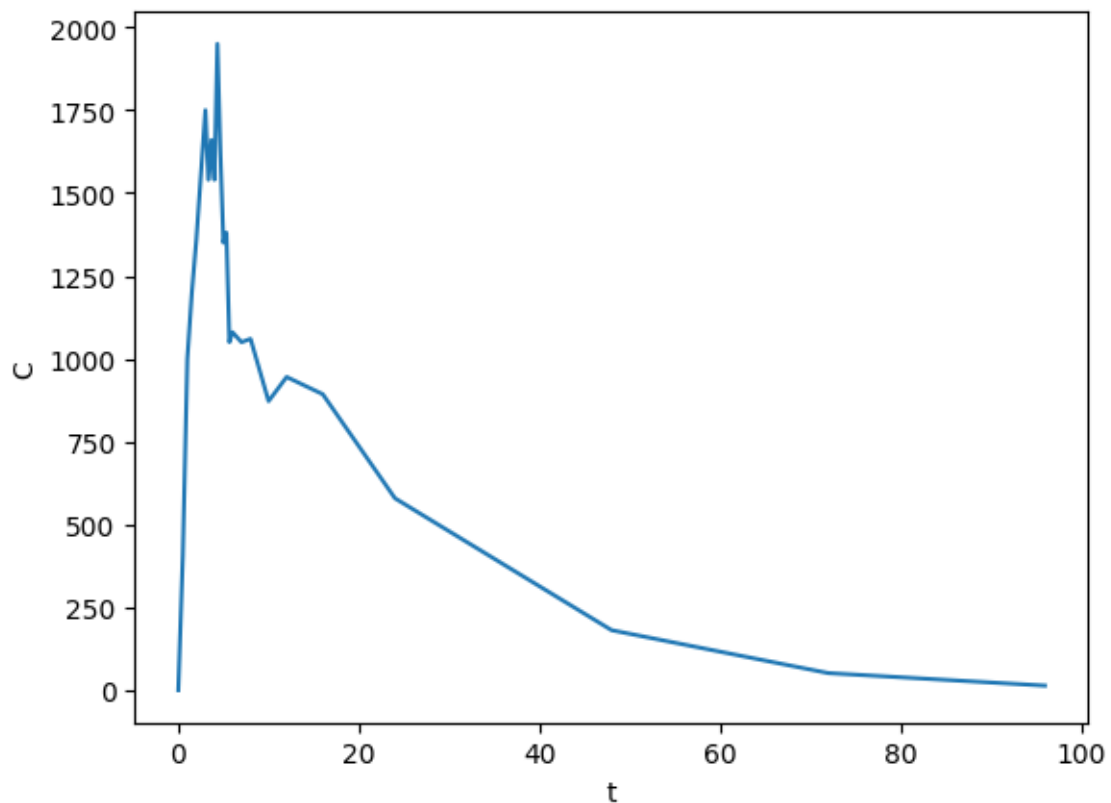
1. Разработать алгоритмы оценки параметров, способные корректно работать при ограниченном количестве данных и высокой степени их зашумлённости.

2. Провести аналитическое и численное исследование устойчивости полученных решений.
3. Проверить предложенную модель на различных наборах реальных и синтетических данных, сравнив результаты с классическими подходами.

Данная постановка задачи позволит более точно описывать фармакокинетические процессы в условиях, близких к реальным, и предоставит исследователям удобный инструмент для интерпретации и прогнозирования эффективных режимов лекарственной терапии.

3 Описание данных

Данные для исследований были предоставлены Центром научного консультирования. Несмотря на то, что данные являются синтетическими, в дальнейших исследованиях будем ссылаться на них как на реальные, поскольку они крайне близки к данным, полученным после одного из проведённых тестов биоэквивалентности двух реальных препаратов. Результаты представлены в виде временных рядов - зависимостей концентрации вещества в крови от времени. Каждый из временных рядов относится к одному из препаратов - тестовому либо референтному. Рассмотрим типичную траекторию имеющегося процесса:



Следует описать основные особенности имеющихся процессов:

- Траектория стартует из нуля.
- Данные неоднородны по времени. Хорошо видно, что в самом начале гораздо больше замеров нежели чем ближе к концу траектории.

- Единицы измерения по разным осям имеют разные порядки, то есть данные не нормированы.
- Сначала концентрация лекарства растёт, затем (с некоторыми оговорками) она убывает.

Все дальнейшие рассуждения, связанные с моделями, будут учитывать вышеуказанные особенности.

4 Описание моделей

4.1 PBFTPК

Существуют две различные модификации **PBFTPК**-модели — **PBFTPК₀** и **PBFTPК₁**. Для их описания введём следующие обозначения:

- $C(t)$ - концентрация лекарства в крови
- F - биодоступная доля лекарственного средства
- D - введённая доза лекарственного средства
- V_d - объём распределения лекарственного средства
- k_a - коэффициент всасываемости лекарственного средства
- k_{el} - коэффициент выводимости лекарственного средства
- τ - время абсорбции лекарственного средства

Также условимся называть $C(t)$ при $t < \tau$ ($t \geq \tau$) этапом абсорбции (выведения)

PBFTPК₀:

$$C(t) = \begin{cases} \frac{FD}{\tau} \frac{1}{V_d k_{el}} (1 - e^{-k_{el}t}), & t \leq \tau \\ C(\tau) e^{-k_{el}(t-\tau)}, & t > \tau \end{cases}$$

PBFTPК₁:

$$C(t) = \begin{cases} \frac{FDk_a}{V_d(k_a - k_{el})} (e^{-k_{el}t} - e^{-k_a t}), & t \leq \tau \\ C(\tau) e^{-k_{el}(t-\tau)}, & t > \tau \end{cases}$$

Недостатком таких моделей является то, что они описывают исключительно процессы с однопиковыми траекториями. Если же траектория процесса может иметь более одного пика, такая модель становится гораздо менее точной. Для исправления этого недостатка можно применить технику ансамблирования моделей.

4.2 Кусочная модификация РВФТРК

Модифицируем уже имеющиеся методы оценки параметров модели. Введём параметры τ_0 и τ_{max} и определим новую модель для случая **РВФТРК₁** (для случая **РВФТРК₀** аналогично):

$$C(t) = \begin{cases} 0, t \leq \tau_0 \\ \frac{FDk_a}{V_d(k_a - k_{el})}(e^{-k_{el}t} - e^{-k_a t}), \tau_0 < t \leq \tau \\ C(\tau)e^{-k_{el}(t-\tau)}, t > \tau \end{cases}$$
$$\tau_0 < \tau < \tau_{max}$$

Параметр τ_{max} является гиперпараметром и задаётся до начала оценки остальных параметров. Для поиска τ_0 и τ существуют два метода: минимаксный и пиковый. Рассмотрим оба этих метода подробнее.

Минимаксный метод

При использовании этого метода оценки выглядят следующим образом:

$$\tau = \arg \max_{t < \tau_{max}} C(t)$$

$$\tau_0 = \arg \min_{t < \tau} C(t)$$

Этот метод особенно хорошо себя показывает в тех ситуациях, когда траектория имеет ровно одну точку максимума и одну точку минимума, удовлетворяющие необходимым условиям. Если же траектория имеет более сложную форму, более полезным будет второй метод.

Пиковый метод

$$\tau = \max_{t < \tau_{max}} \text{Arg max } C(t)$$

$$\tau_0 = \max_{t < \tau} \text{Arg min } C(t)$$

То есть, вместо поиска условных максимума и минимума функции этот метод выбирает «максимальные» точки максимума и минимума соответственно. Основное преимущество этого метода заключается в том, что остатки $r(t) = C(t) - \hat{C}(t)$ будут иметь структуру похожую на исходную траекторию, что пригодится в дальнейшем.

4.3 ЕРВФТРК

Возьмём в качестве базовой модели описанную выше модификацию **PВФТРК**. Примем $\tau_{max}^1 = +\infty$. Далее одним из перечисленных выше методов оценим параметры τ^1 и τ_0^1 : $\hat{\tau}^1$, $\hat{\tau}_0^1$. После поиска параметров **PВФТРК** получим некую оценку концентрации $\hat{C}_1(t)$. Рассмотрим полученные остатки $r(t) = C(t) - \hat{C}(t)$. Далее тем же методом построим оценку $\hat{C}_2(t)$ для остатков, причём в качестве параметра τ_{max}^2 возьмём полученную ранее оценку $\hat{\tau}^1$. Будем повторять эти шаги до тех пор, пока будет выполняться некий заранее установленный критерий эффективности модели. Итоговая оценка траектории $\hat{C}(t)$ тогда получается следующим образом:

$$\hat{C}(t) = \sum_i \hat{C}_i(t)$$

Полученную модель будем называть **ЕРВФТРК** (ensembled PВФТРК).

4.4 Функции потерь

Для работы с полученными моделями будем использовать следующую функцию потерь:

$$L_\lambda = \frac{1}{N} \sum_{i=1}^N l_\lambda(C, \hat{C}, t_i)$$

где

$$l_\lambda(C, \hat{C}, t) = (C(t) - \hat{C}(t))^2 \cdot \begin{cases} 1, t \leq \tau \\ \lambda, t > \tau \end{cases}$$

Здесь и далее будем называть эту функцию λ - взвешенной среднеквадратичной функцией потерь (**WMSE** $_\lambda$). Основным достоинством такой функции является возможность регулировать влияние точности на этапе абсорбции на общую точность модели. При $\lambda < 1$ более значительной будет считаться ошибка оценки на этапе абсорбции, при $\lambda > 1$ - на этапе выведения. При $\lambda = 1$ данная функция является обыкновенной квадратичной функцией потерь (**MSE**).

Стоит заметить, что эта функция в дальнейшем поможет нивелировать недостаток данных, связанный с более редкими измерениями после времени абсорбции.

4.5 Описание итогового алгоритма

В итоге, задача оценки траектории с помощью модели сводится к задаче последовательной минимизации:

$$L_\lambda(r_i, \hat{C}_i, t) \rightarrow \inf_{p \in \Omega}$$

где

$$\Omega = \{(F, D, V_d, k_a, k_{el}) | F \in [0; 1], D \geq 0, V_d \geq 0, k_a \geq 0, k_{el} \geq 0\}$$

$$r_1 = C$$

Также стоит упомянуть, что данный метод при незначительной модификации может быть использован для предсказания траектории. Для этого нужно будет изменить l_λ таким образом, что $l_\lambda = 0$ при $t > \tau_{max}$, и считать за τ_{max} время последнего замера концентрации лекарства в крови.

5 Полученные результаты

На графиках представлены сравнительные результаты работы двух моделей - **РВФТРК** (сверху) и полученной в ходе исследований модели **ЕРВФТРК** с взвешенной среднеквадратичной функцией потерь (снизу). Можно заметить, что полученная модель при удачном подборе гиперпараметров (количество моделей в ансамбле, коэффициент λ , метод оценки абсорбции, и т. д.) оказывается более точной и лучше отражает зависимости в реальных данных. Особенно хорошо это видно в тех случаях, когда после времени абсорбции концентрация убывает не сразу. Стоит отметить, что наименьшую точность модель будет иметь на «пиках», то есть как раз в моменты τ_i .

Здесь будет таблица со значениями MSE для различных моделей

Рис. 1: РВФТРК (Пример 1)

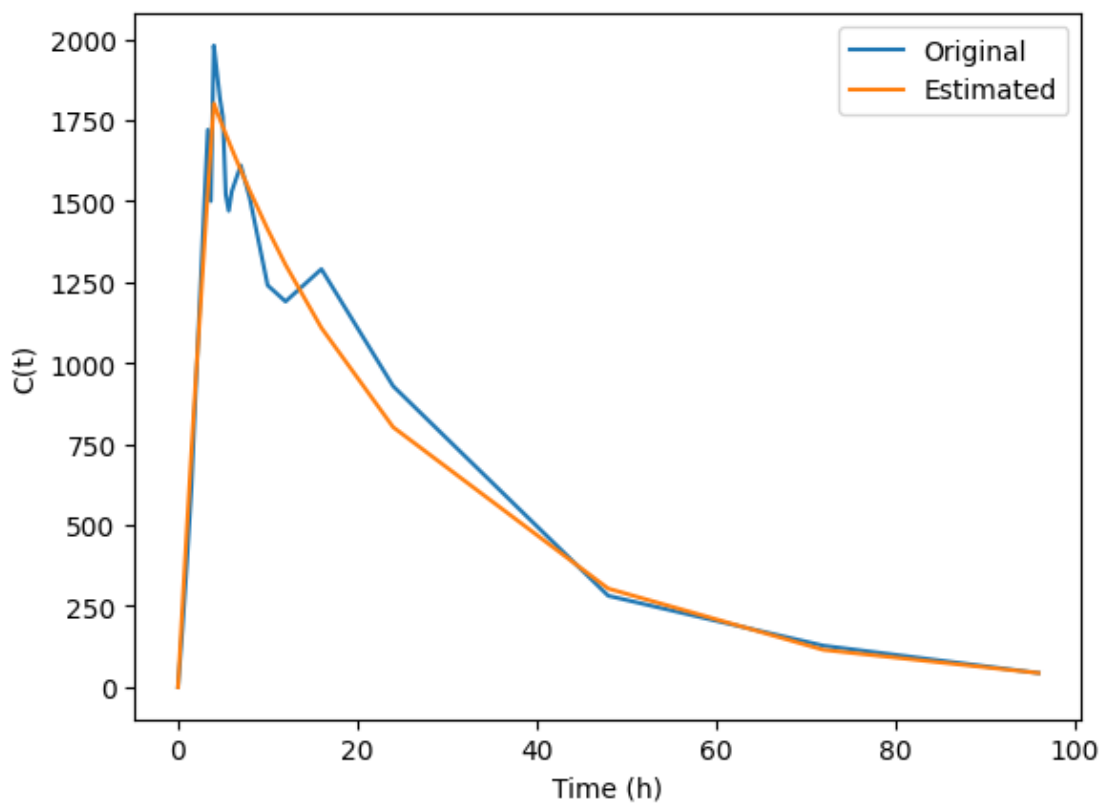


Рис. 2: ЕРВФТРК (Пример 1)

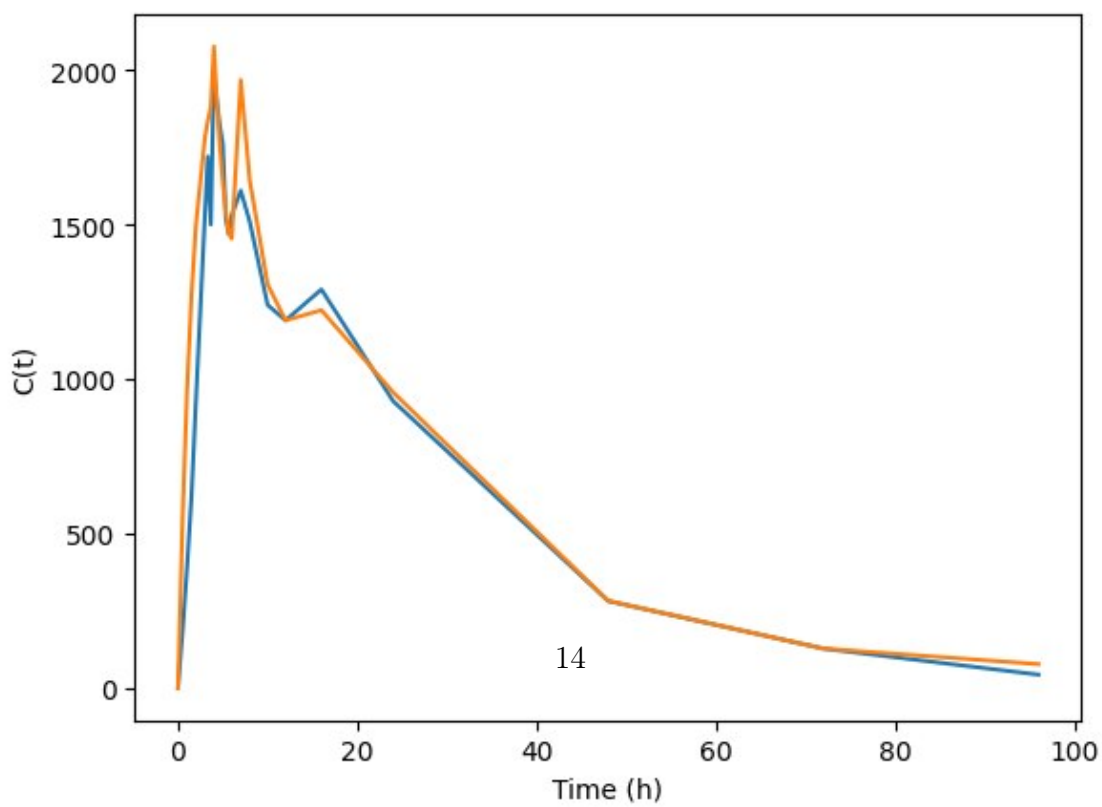


Рис. 3: РВФТРК (Пример 2)

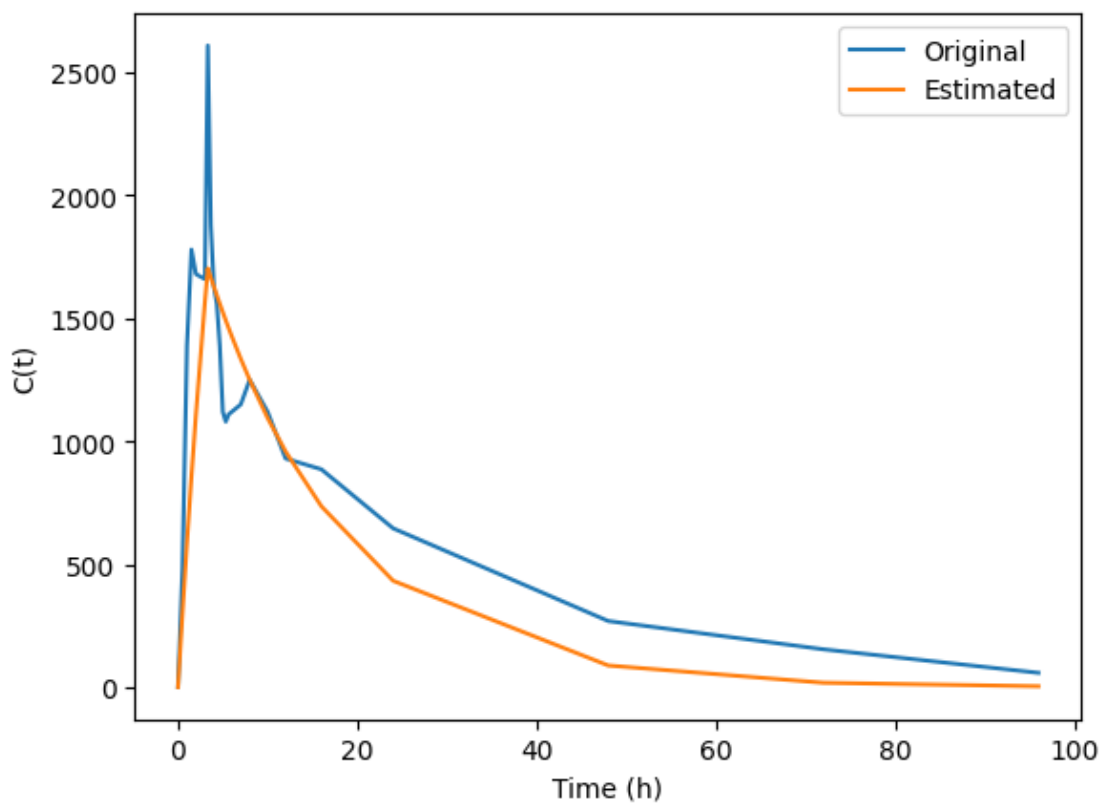


Рис. 4: ЕРВФТРК (Пример 2)

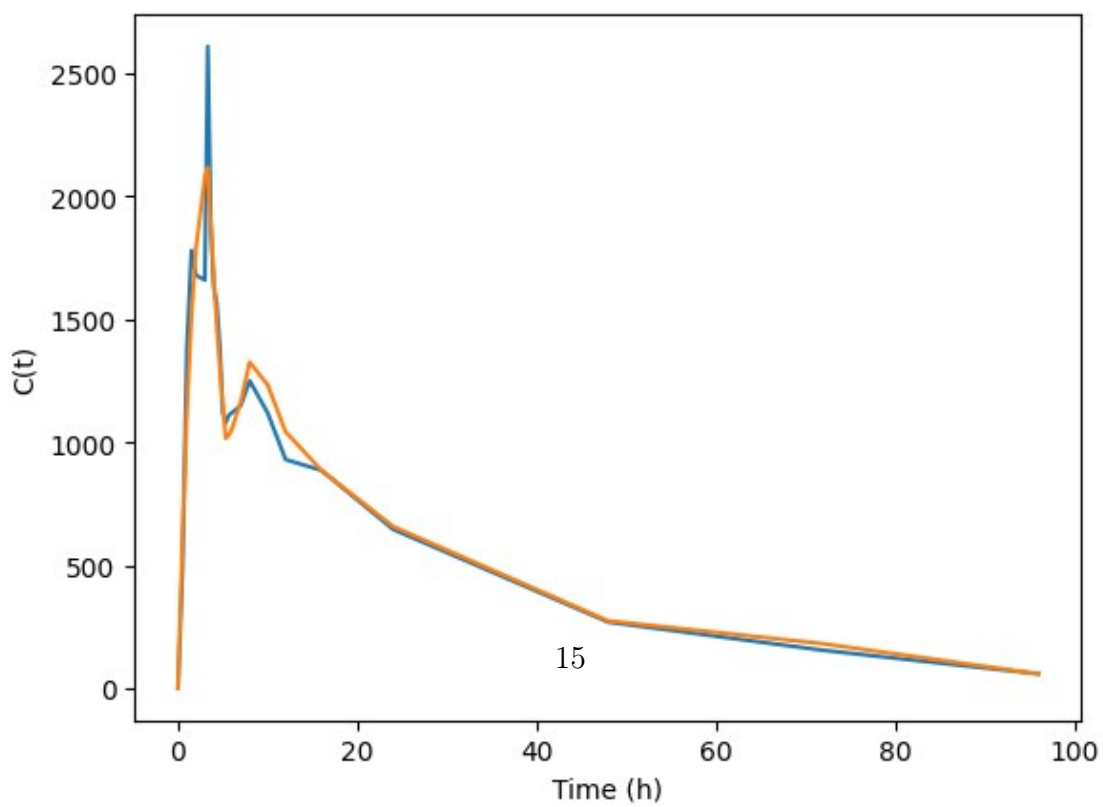


Рис. 5: РВФТРК (Пример 3)

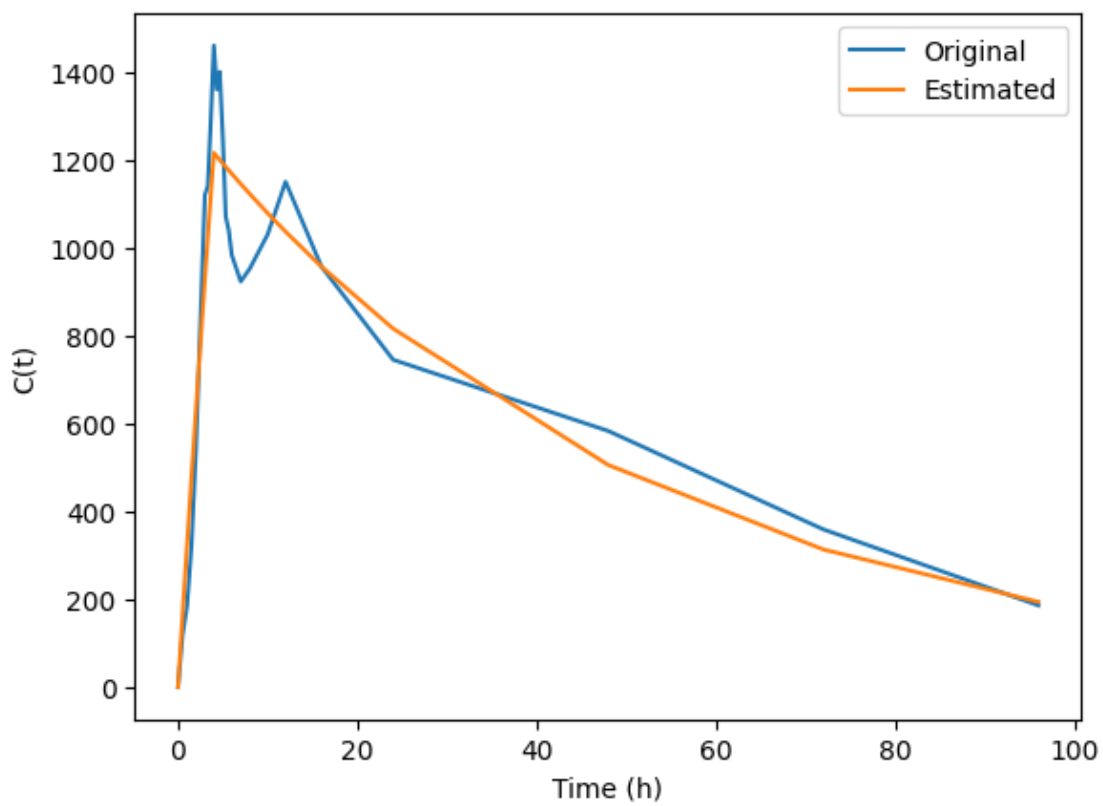


Рис. 6: ЕРВФТРК (Пример 3)

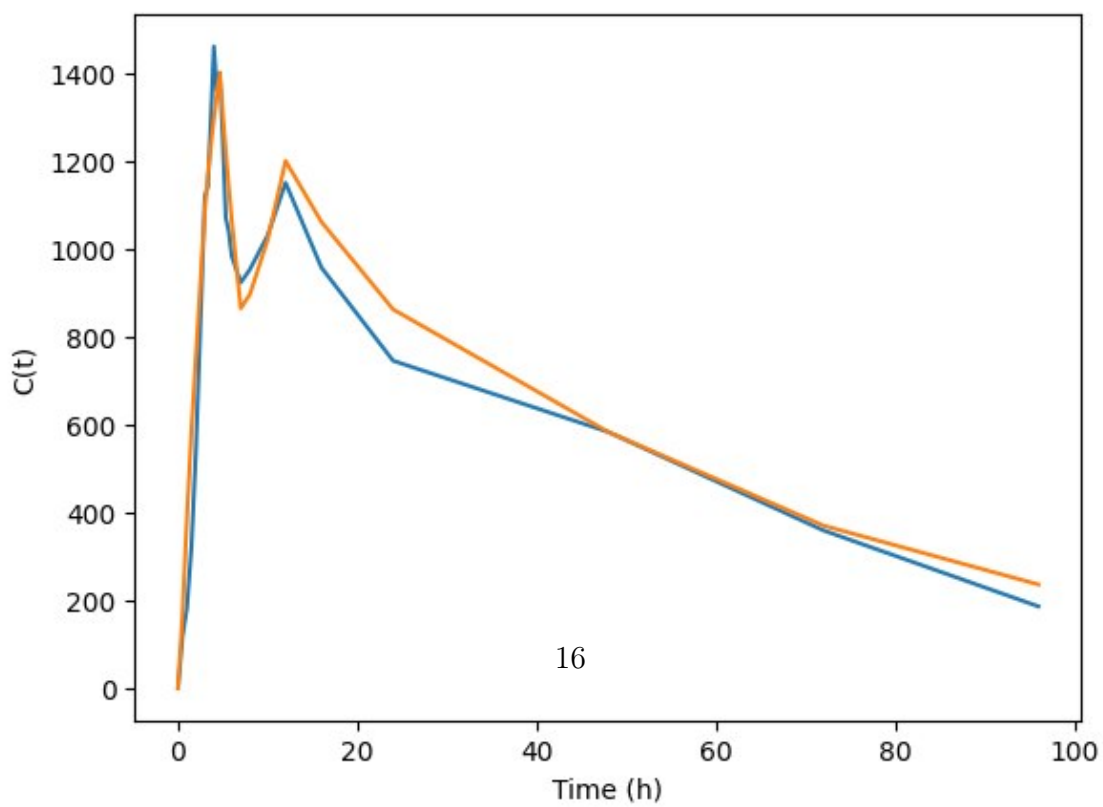


Рис. 7: РВФТРК (Пример 4)

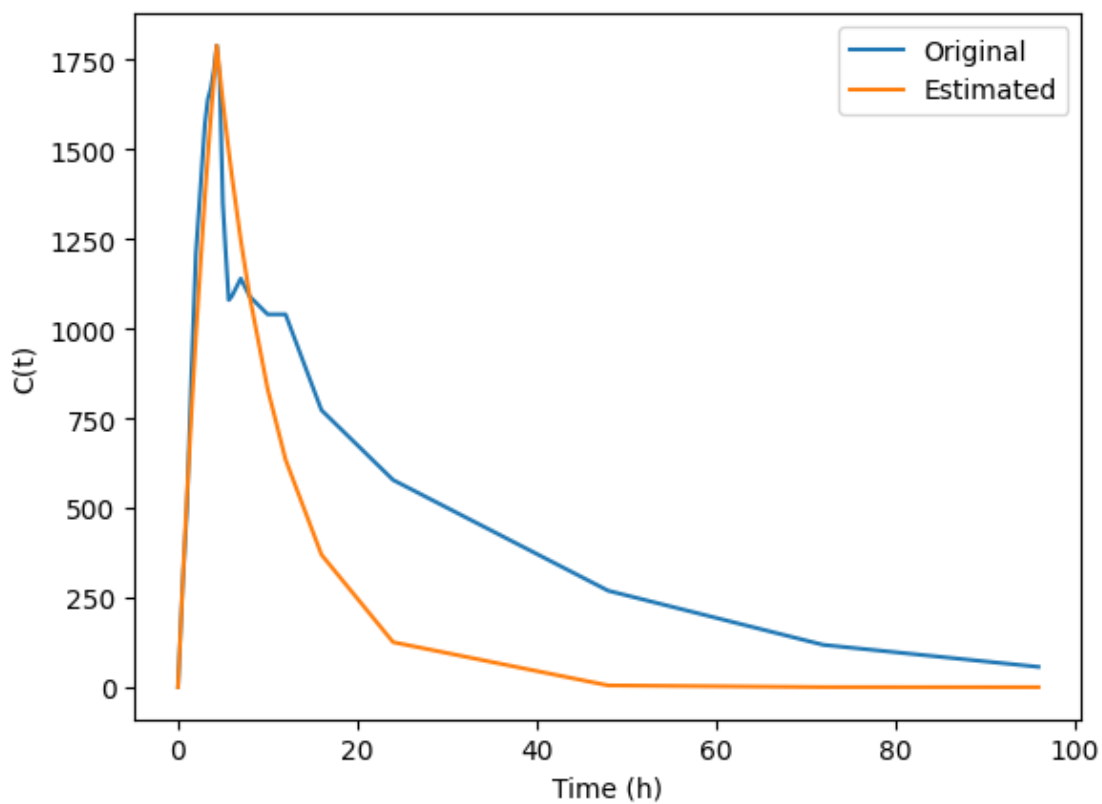


Рис. 8: ЕРВФТРК (Пример 4)

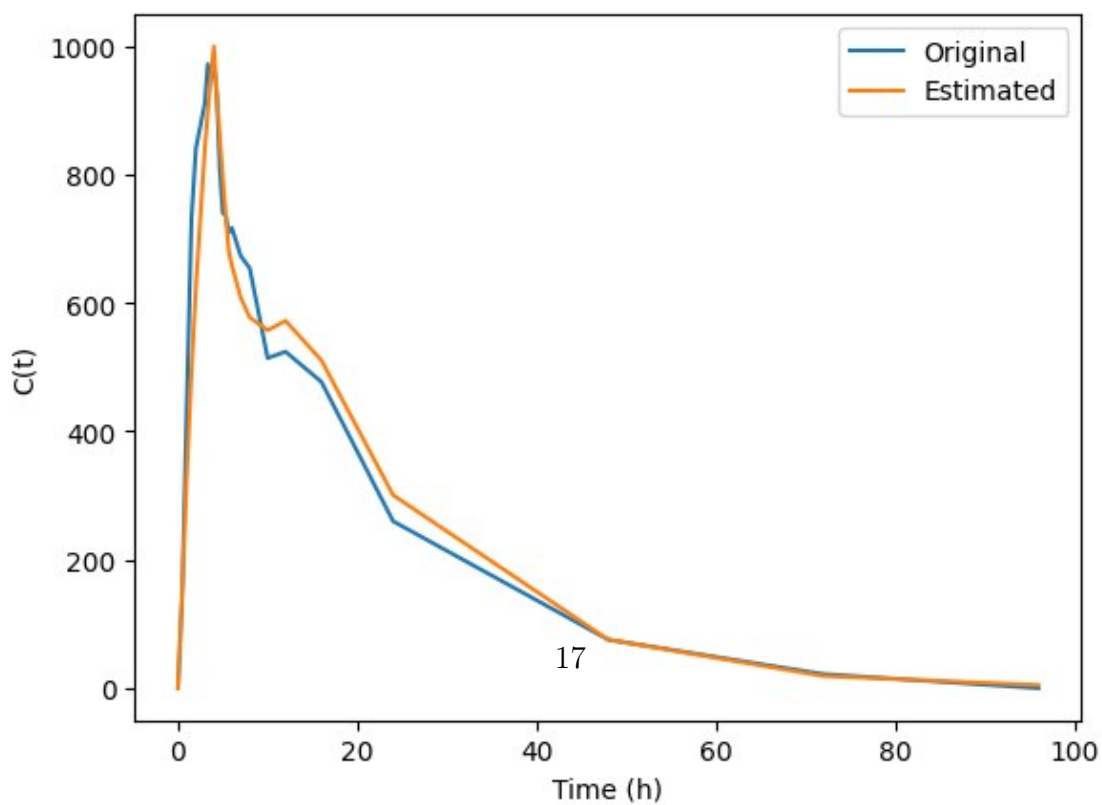


Рис. 9: РВФТРК (Пример 5)

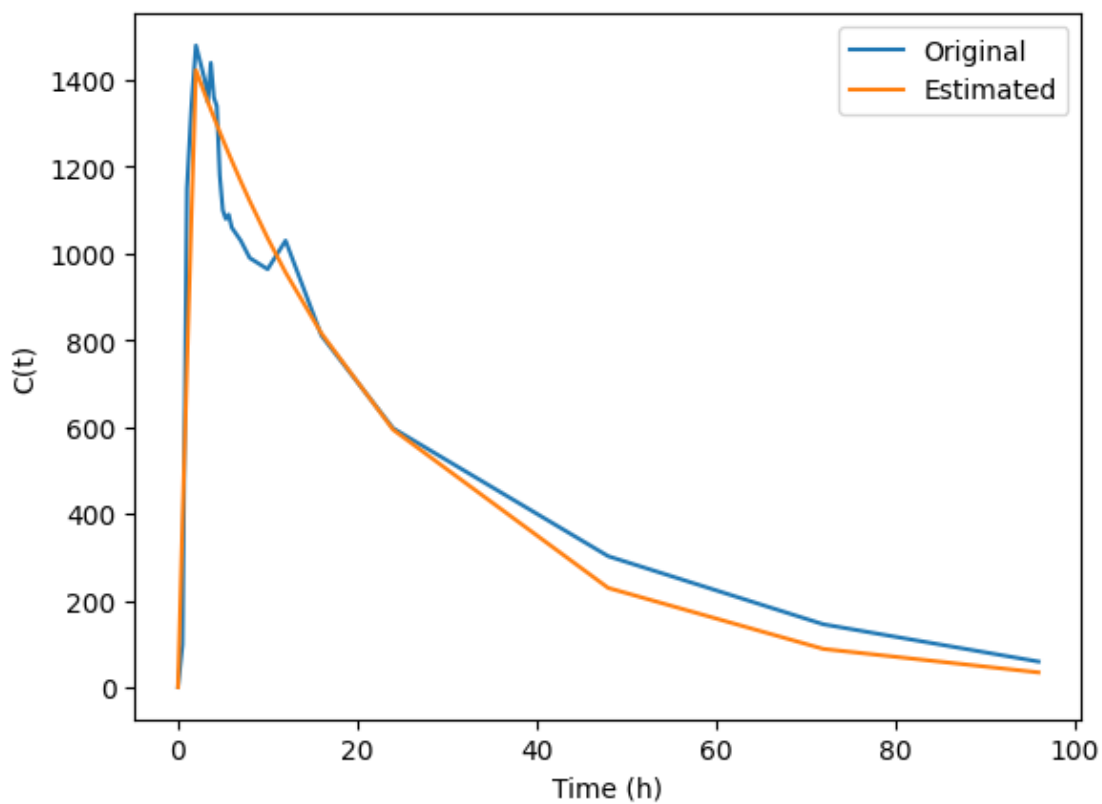


Рис. 10: ЕРВФТРК (Пример 5)

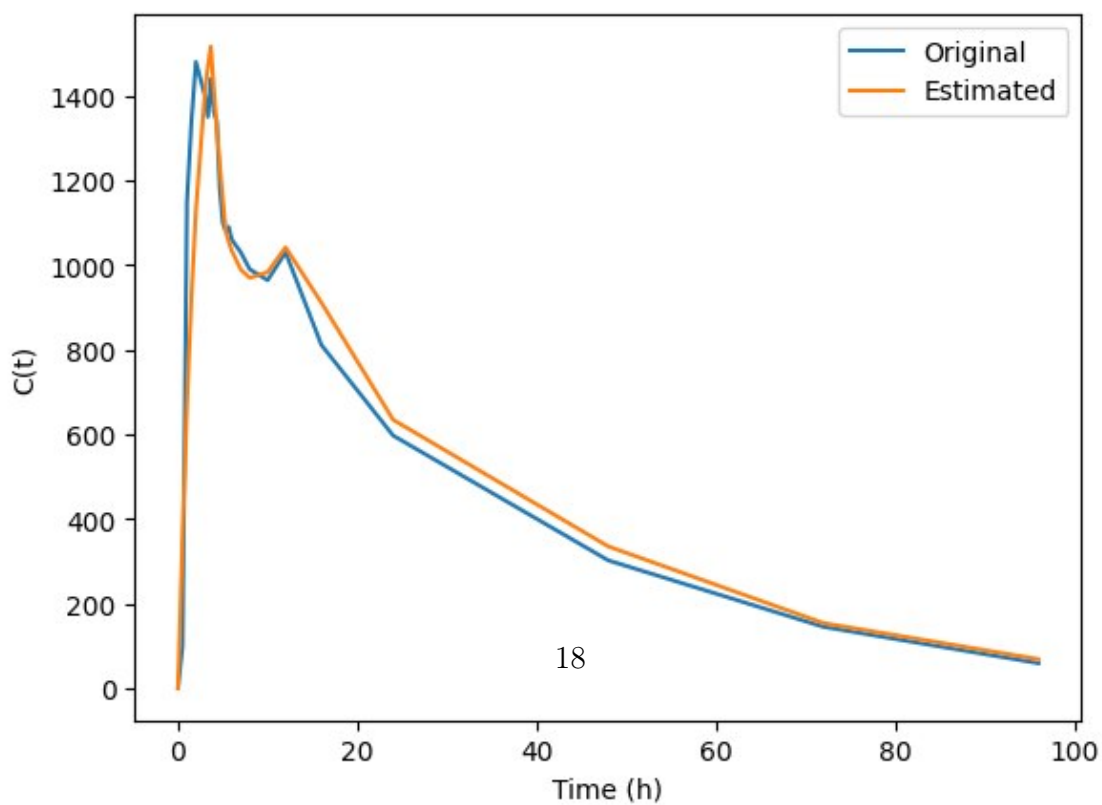


Рис. 11: РВФТРК (Пример 6)

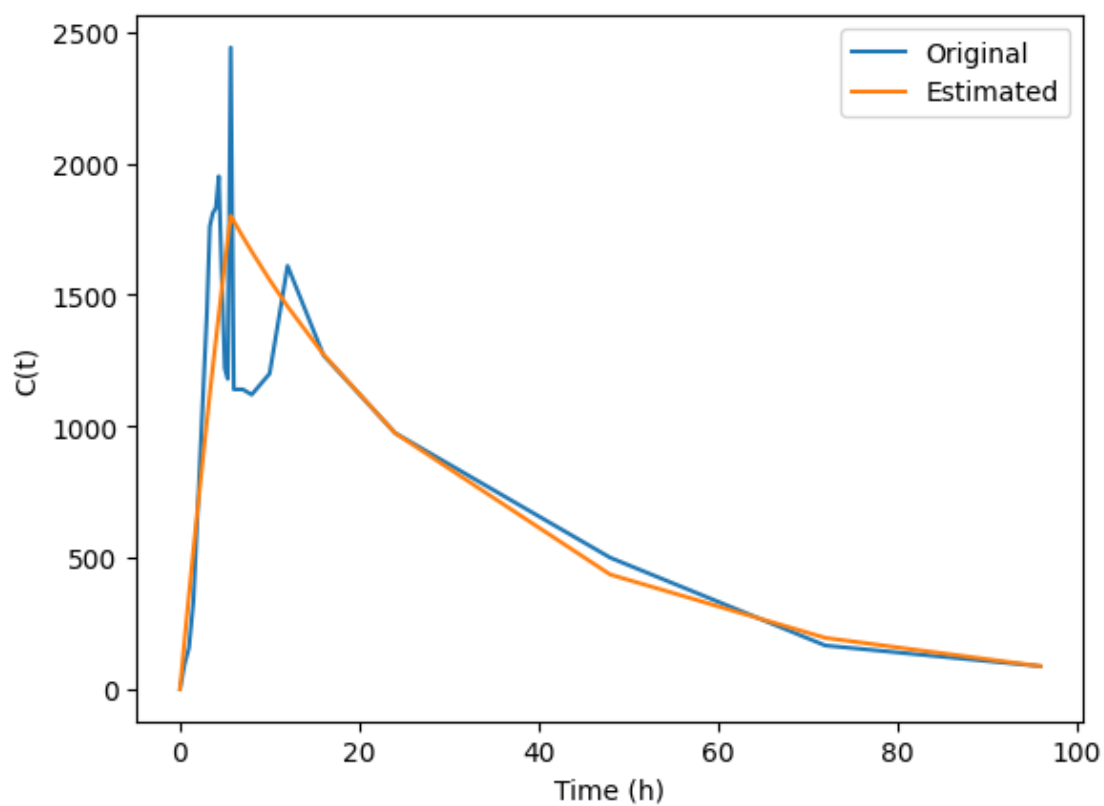
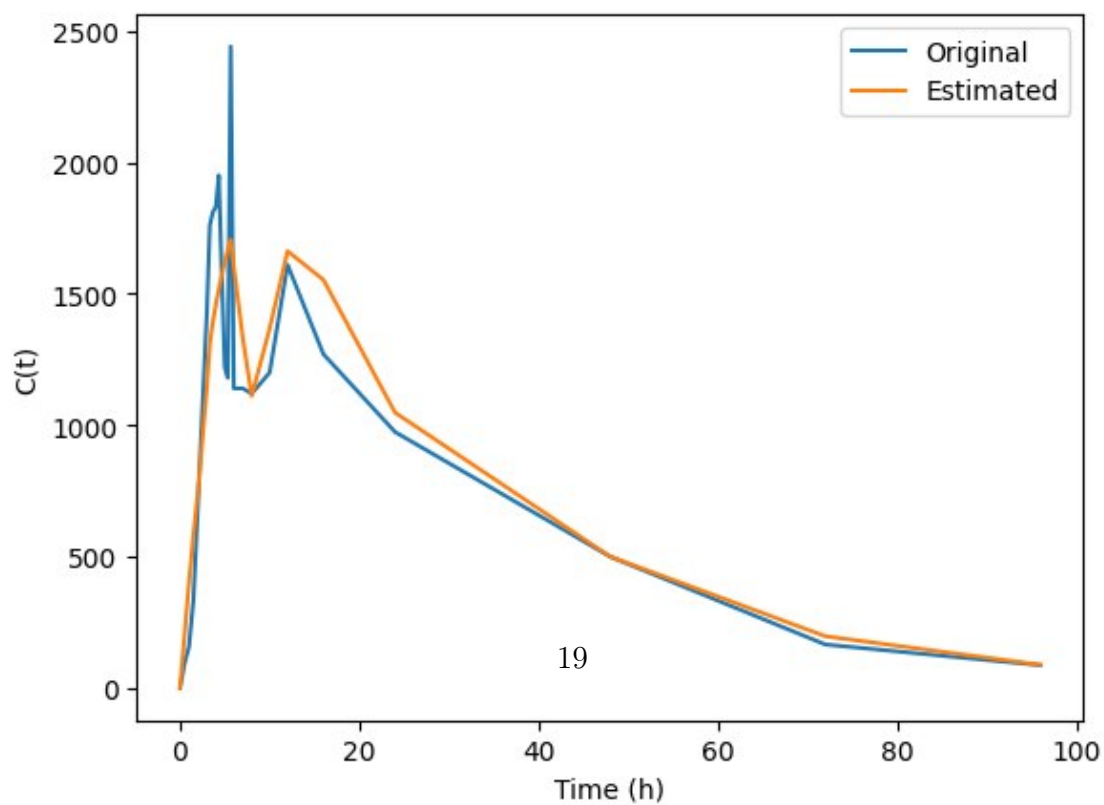


Рис. 12: ЕРВФТРК (Пример 6)



6 Заключение

Бла бла бла.

Список литературы

- [1] *Macheras P., Chrysafidis P.* Revising pharmacokinetics of oral drug absorption: I. Models based on biopharmaceutic, physiological, and finite absorption time concepts. / Chrysafidis P. Macheras P. — 2020. — P. 137.