**به‌نام خدا**

**پدرام رستمی – 9531032**

**گزارش پروژه رگراسیون**

در این پروژه قصد داریم با استفاده از الگوریتم ژنتیک، با یک معادله‌ی خط درجه 3 یک معادله‌ را رگراسیون کنیم. در این مسئله، individual ها همان ضرایب اند. در این گزارش به بررسی تأثیر متغیرهای مسئله‌ی ژنتیک در میزان همگرایی دو معادله خط می‌پردازیم.

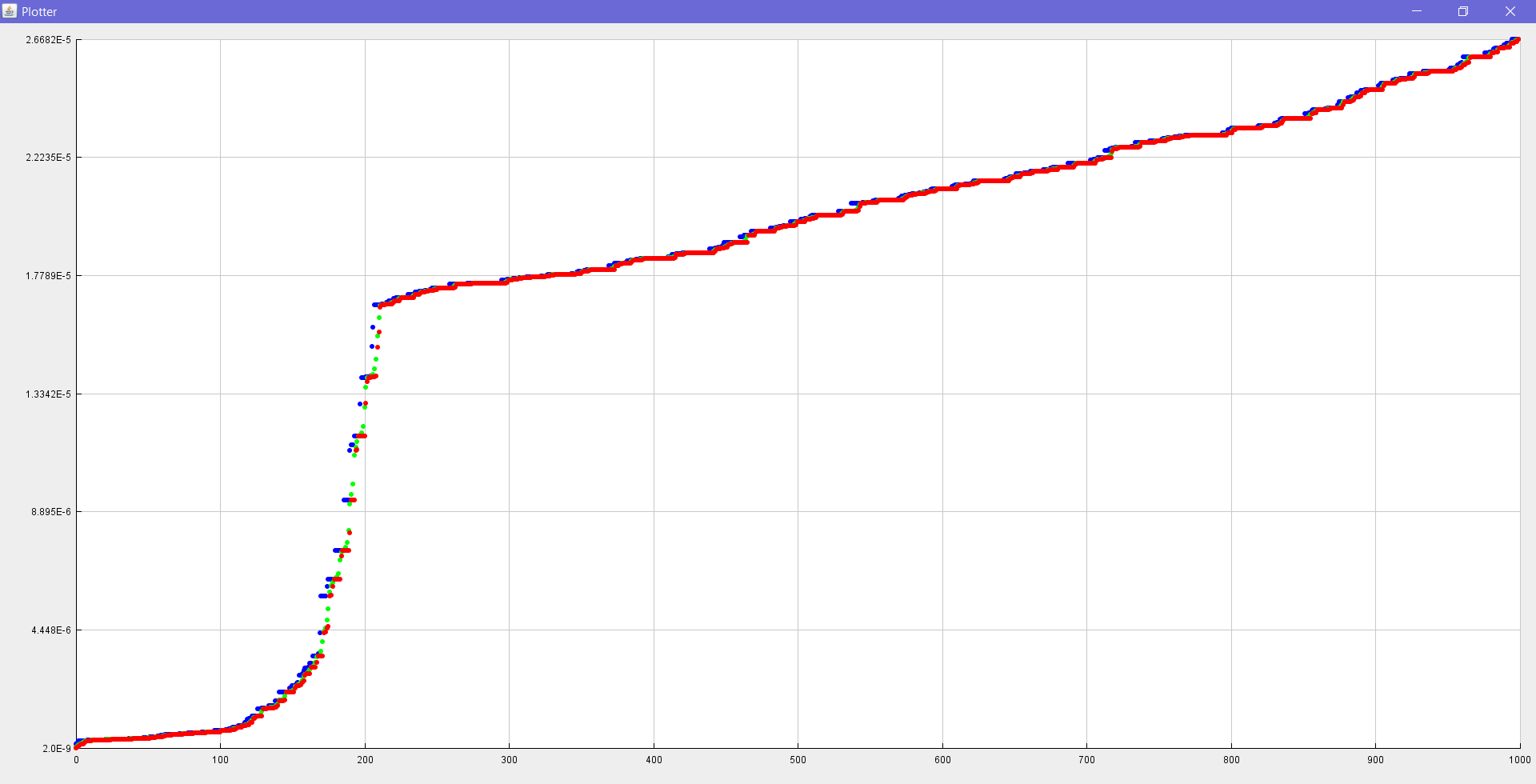
**تأثیر متغیر تعداد نسل‌ها (Number of Generations)**

برای تشخیص شناسایی تأثیر تعداد نسل‌ها بر این مسئله، کافیست متغیر NumberOfGeneration را برابر مقادیر 1000، 2500 و 5000 قرار دهیم و بقیه‌ی متغیرها را ثابت فرض کنیم. برای کاهش تاثیر جمعیت اولیه‌ی تصادفی، برای هر 3 مقدار 1000، 2500 و 5000 از جمعیت اولیه‌ی تصادفی یکسانی استفاده می‌کنیم.

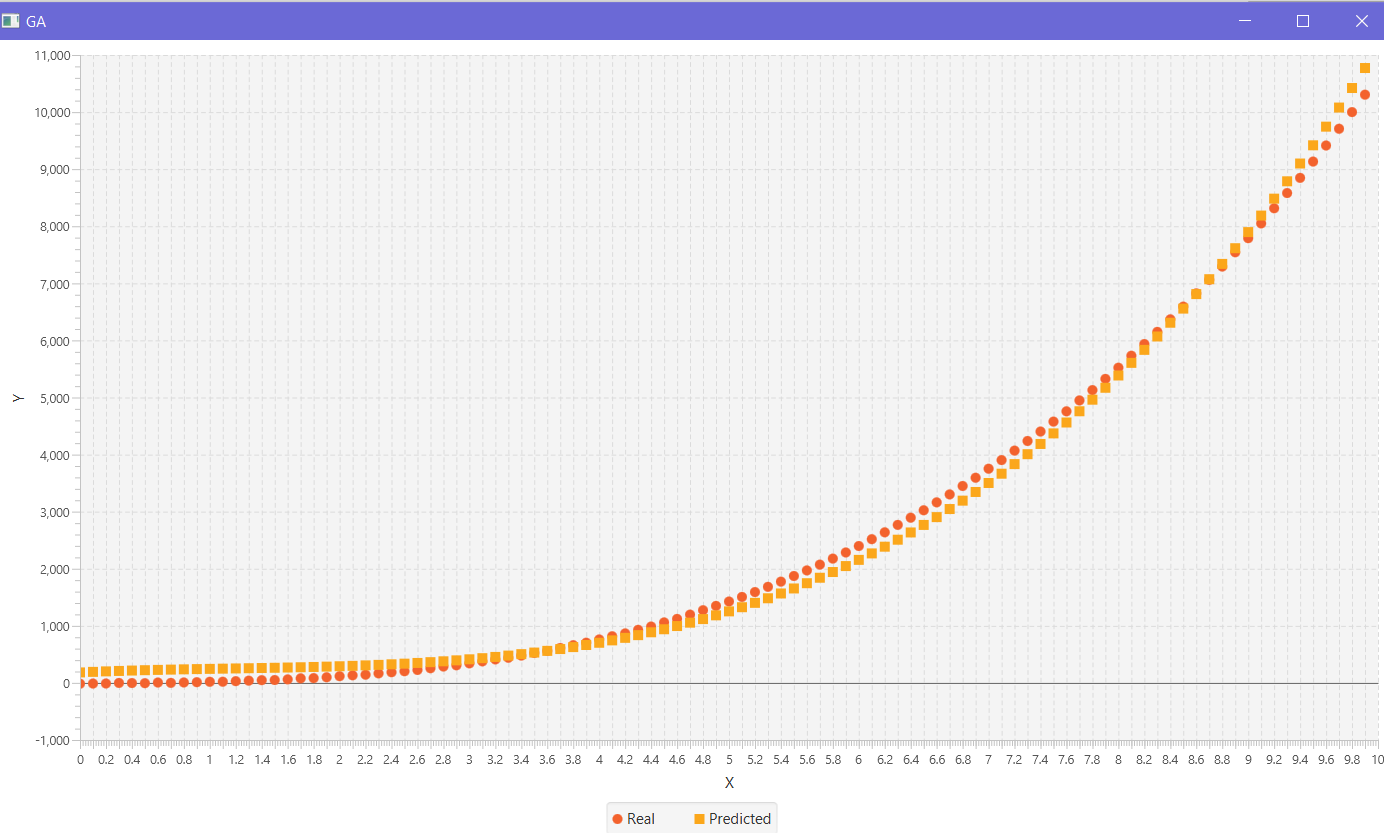
*تست (1) تأثیر تعداد نسل‌ها*

*NumberOfGenerations = 1000 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.02, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -200 تا 200 است.*

**

شکل 1 – نمودار بهترین، بدترین و میانگین شایستگی تست 1



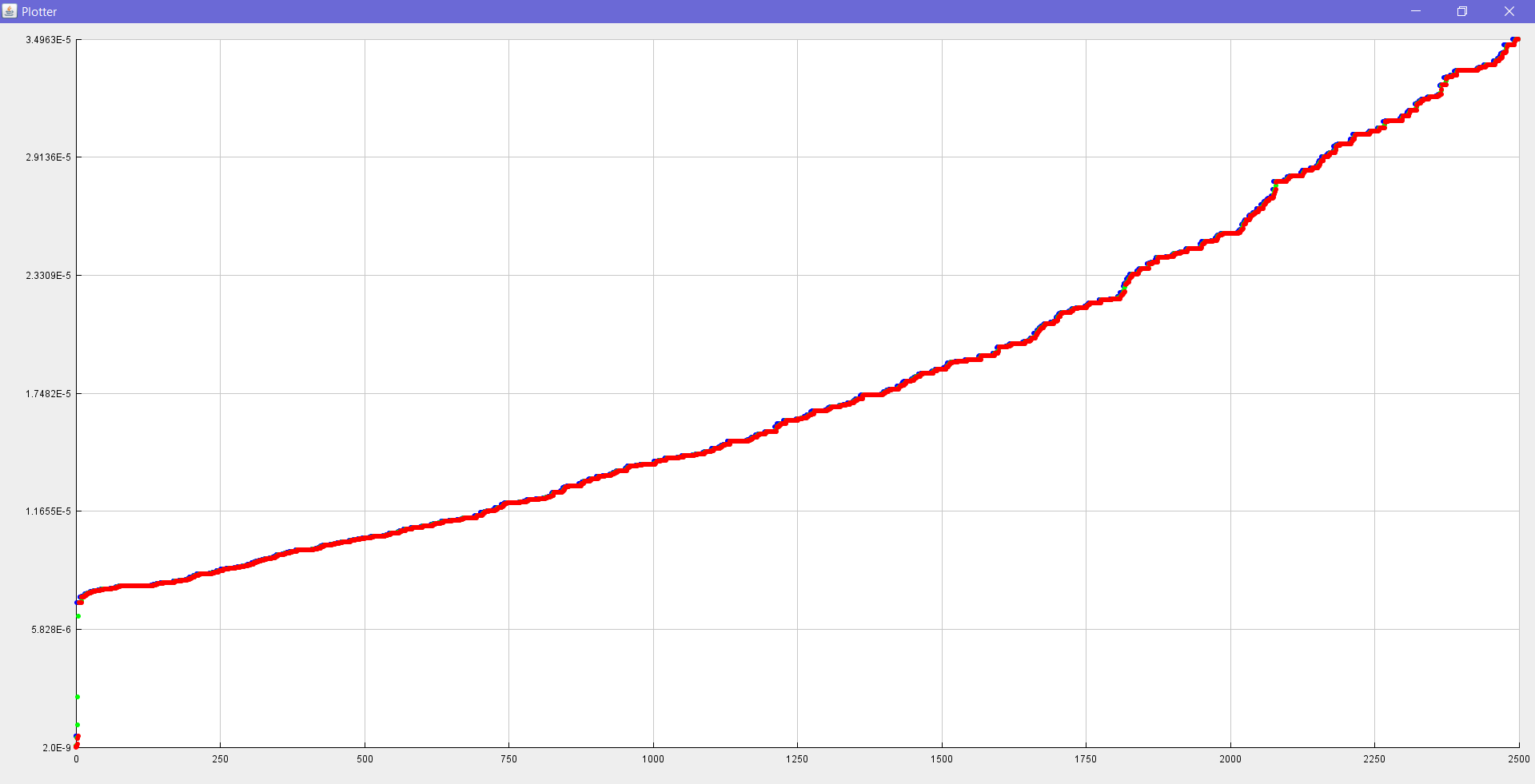
شکل 2 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 1

معادله‌ی پیشبینی‌شده :

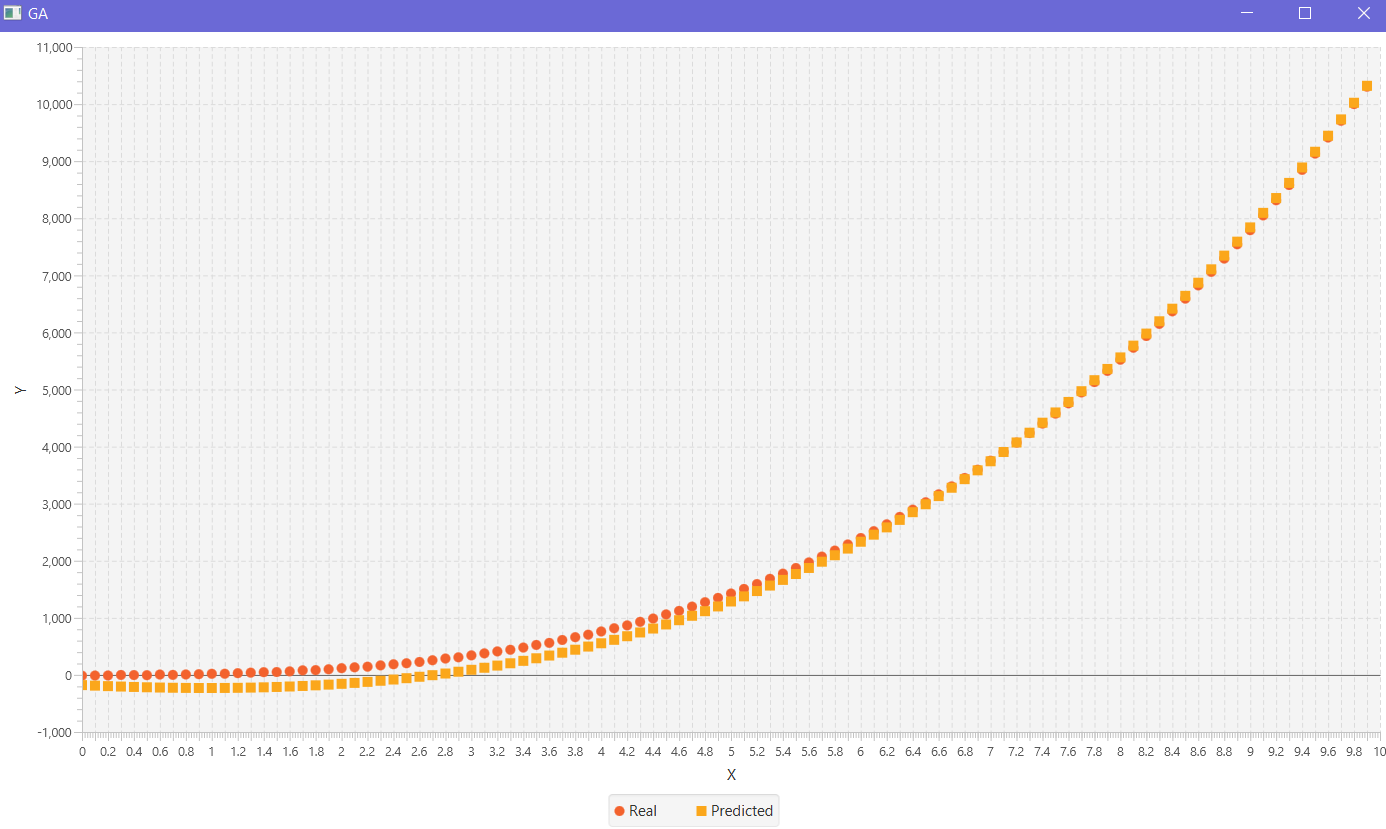
*تست (2) تأثیر تعداد نسل‌ها*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.02, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -200 تا 200 است.*

**

شکل 3 – نمودار بهترین، بدترین و میانگین شایستگی تست 2



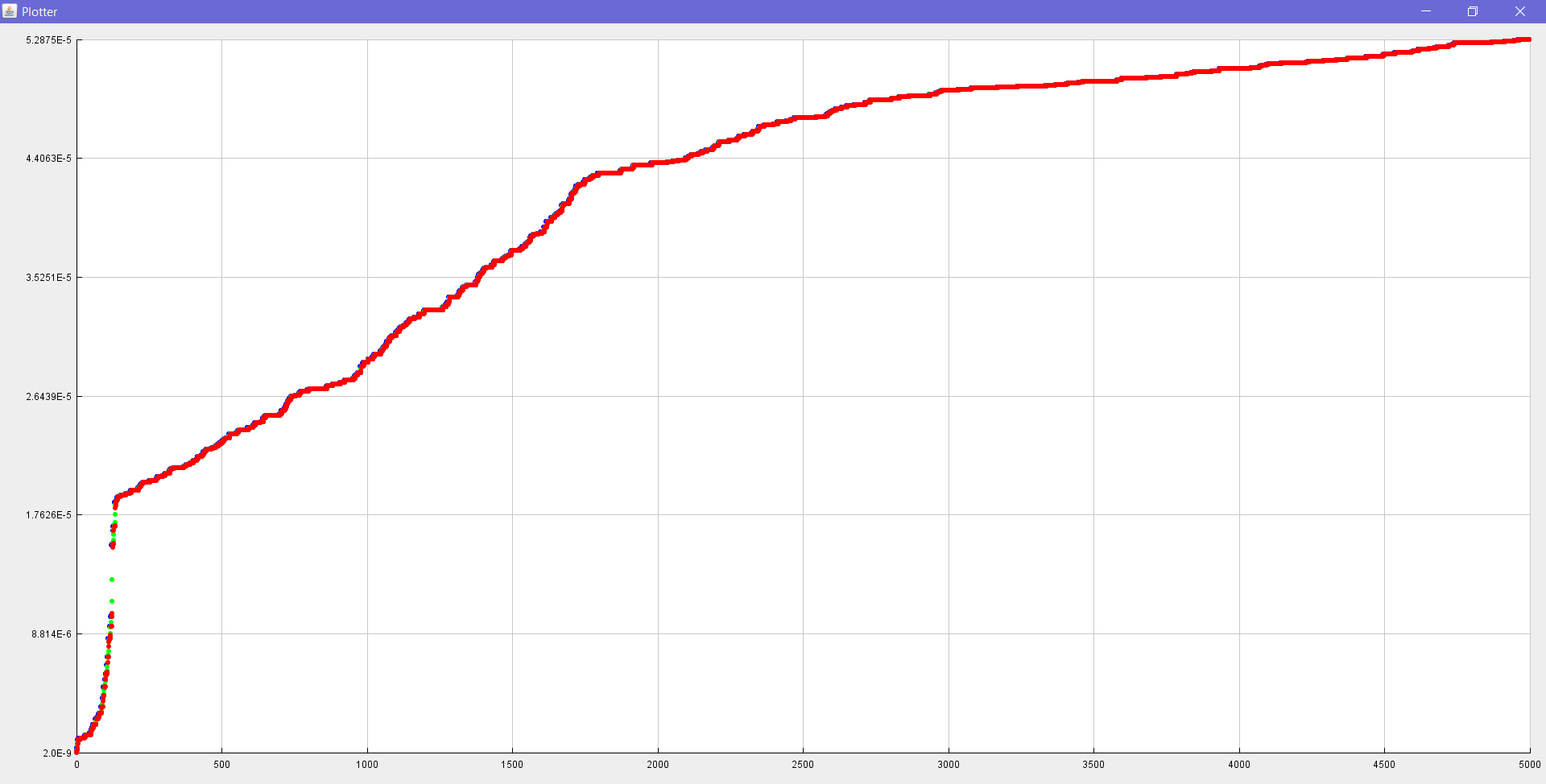
شکل 4 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 2

معادله‌ی پیشینی‌شده:

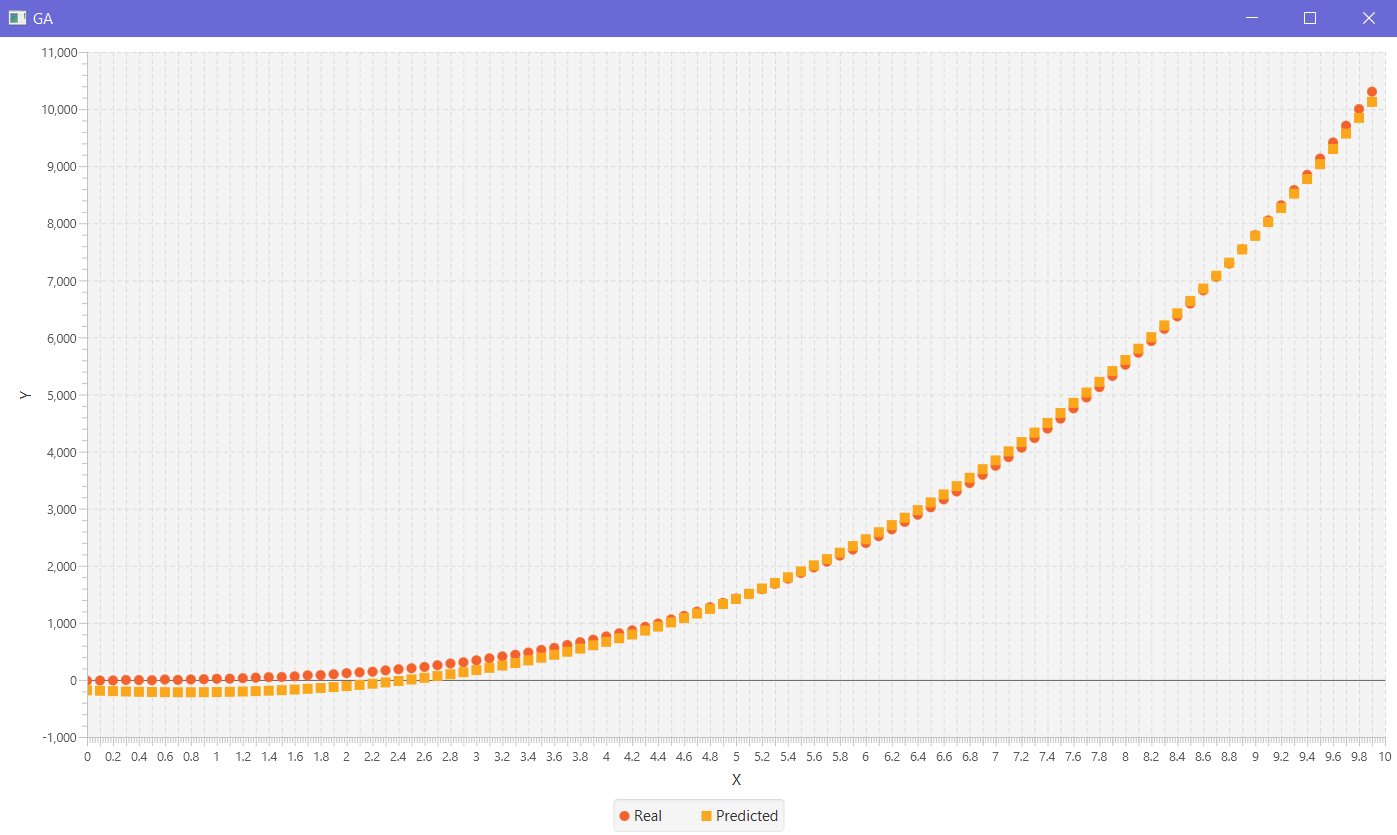
*تست (3) تأثیر تعداد نسل‌ها*

*NumberOfGenerations = 5000 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.02, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -200 تا 200 است.*



شکل 5 – نمودار بهترین، بدترین و میانگین شایستگی تست 3



شکل 6 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 3

معادله‌ی پیشینی‌شده:

**نتیجه‌گیری**

همانطور که در تست‌های (1)، (2) و (3) و جدول 1 مشخص است، با افزایش تعداد نسل‌ها ضرایب بهتری پیدا می‌شوند.

|  |  |
| --- | --- |
| فیتنس آخرین نسل | تعداد نسل‌ها |
| 2.6682E-5 | 1000 |
| 3.4963E-5 | 2500 |
| 5.2875E-5 | 5000 |

جدول 1 – تأثیر تعداد نسل‌ها بر شایستگی آخرین نسل

با افزایش تعداد نسل‌ها، تعداد ژن‌های قویتر در جمعیت بیشتر شده و جمعیت قوی‌تر می‌شود. در نتیجه با افزایش تعداد نسل‌ها، جمعیت شایسته‌تر می‌شود.

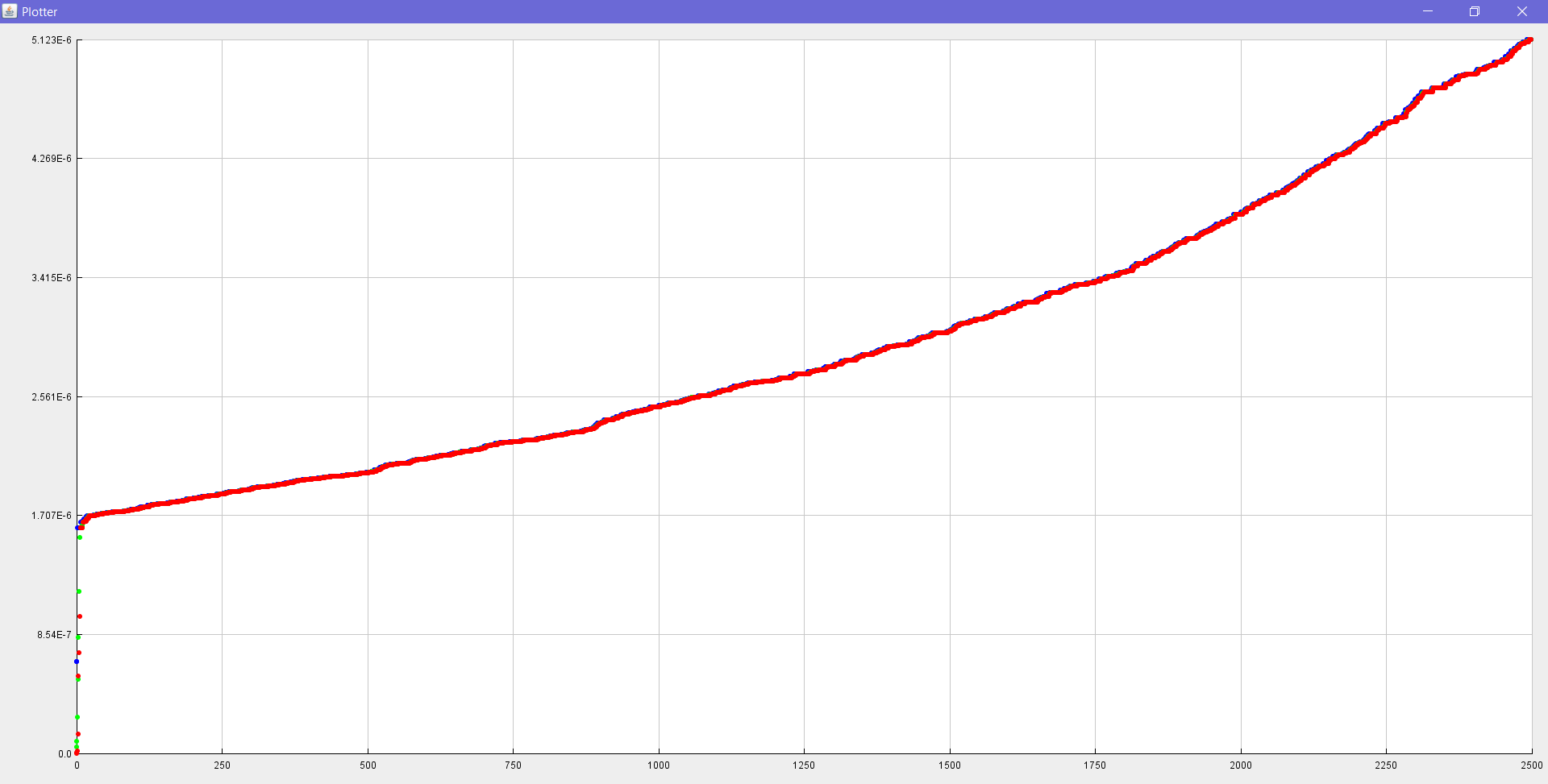
**تأثیر متغیر اندازه‌ی جمعیت (Population Size)**

برای تشخیص شناسایی تأثیر اندازه‌ی جمعیت، کافیست متغیر PopulationSize را برابر مقادیر 50، 100 و 200 قرار می‌دهیم و بقیه‌ی متغیرها را ثابت فرض می‌کنیم.

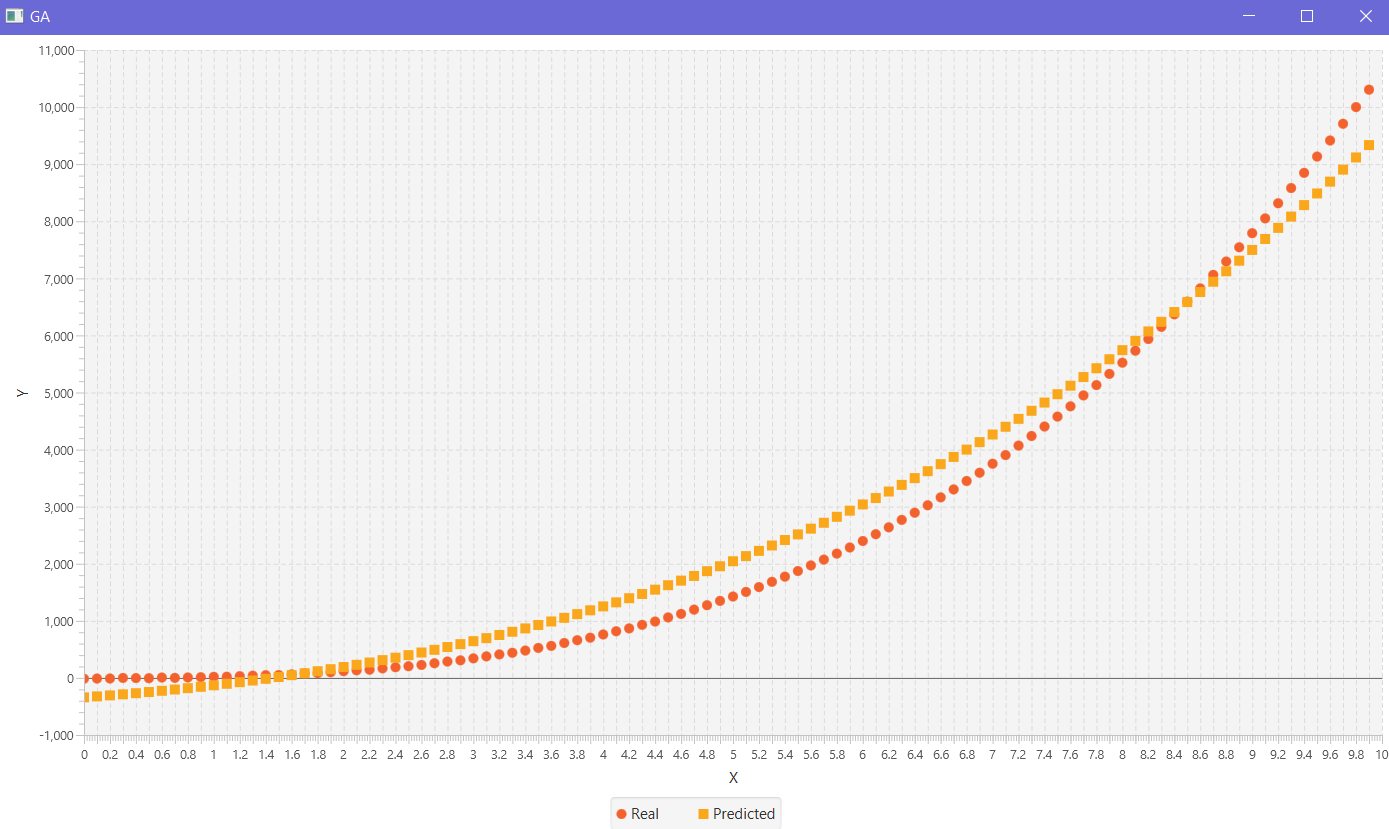
*تست (4) تأثیر اندازه‌ی جمعیت*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.02, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*

**

شکل 7 – نمودار بهترین، بدترین و میانگین شایستگی تست 4



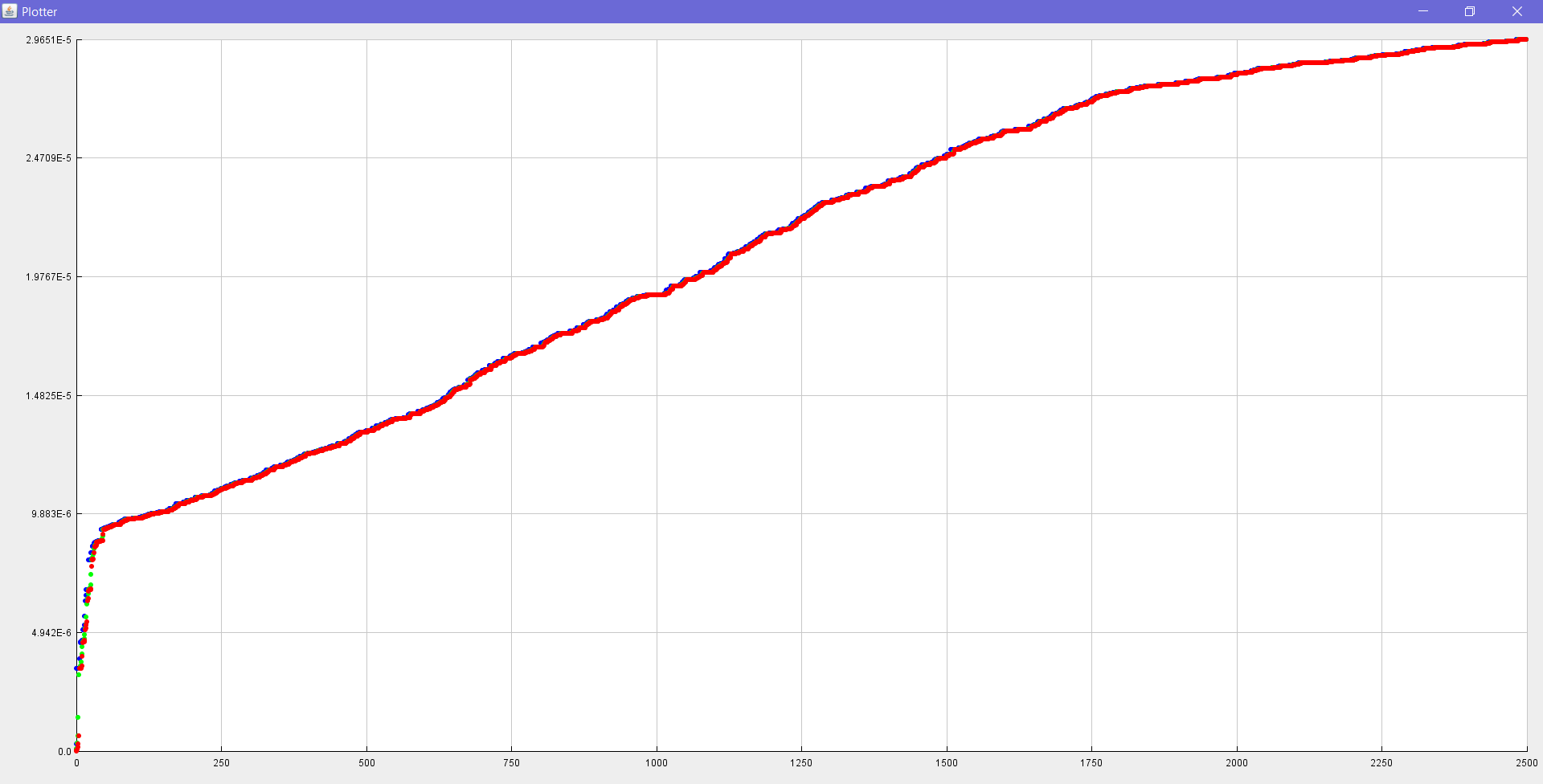
شکل 8 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 4

معادله‌ی پیشینی‌شده:

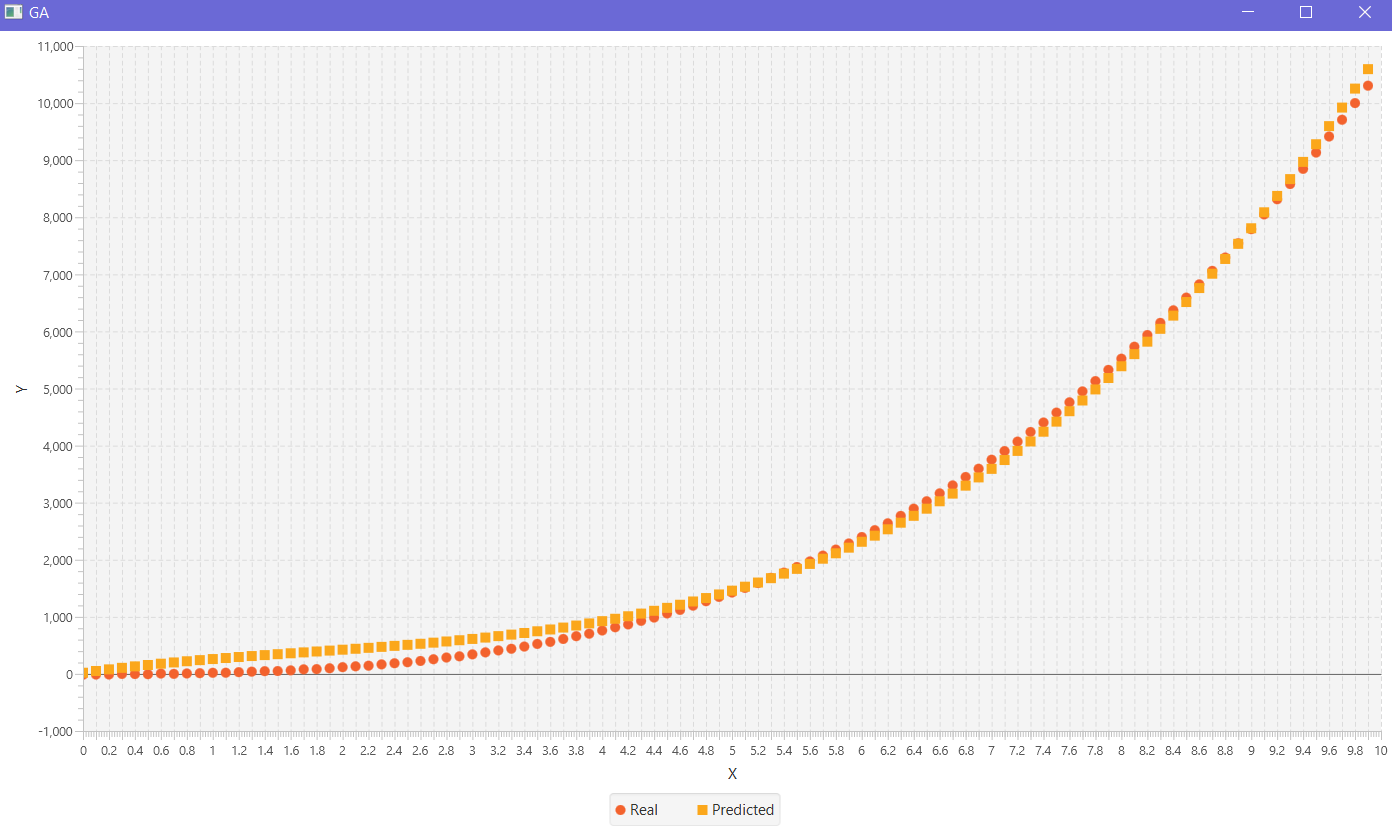
*تست (5) تأثیر اندازه‌ی جمعیت*

*NumberOfGenerations = 2500 , PopulationSize = 100, TournamentSize = 2, MutationRate = 0.02, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*

**

شکل 9 – نمودار بهترین، بدترین و میانگین شایستگی تست 5



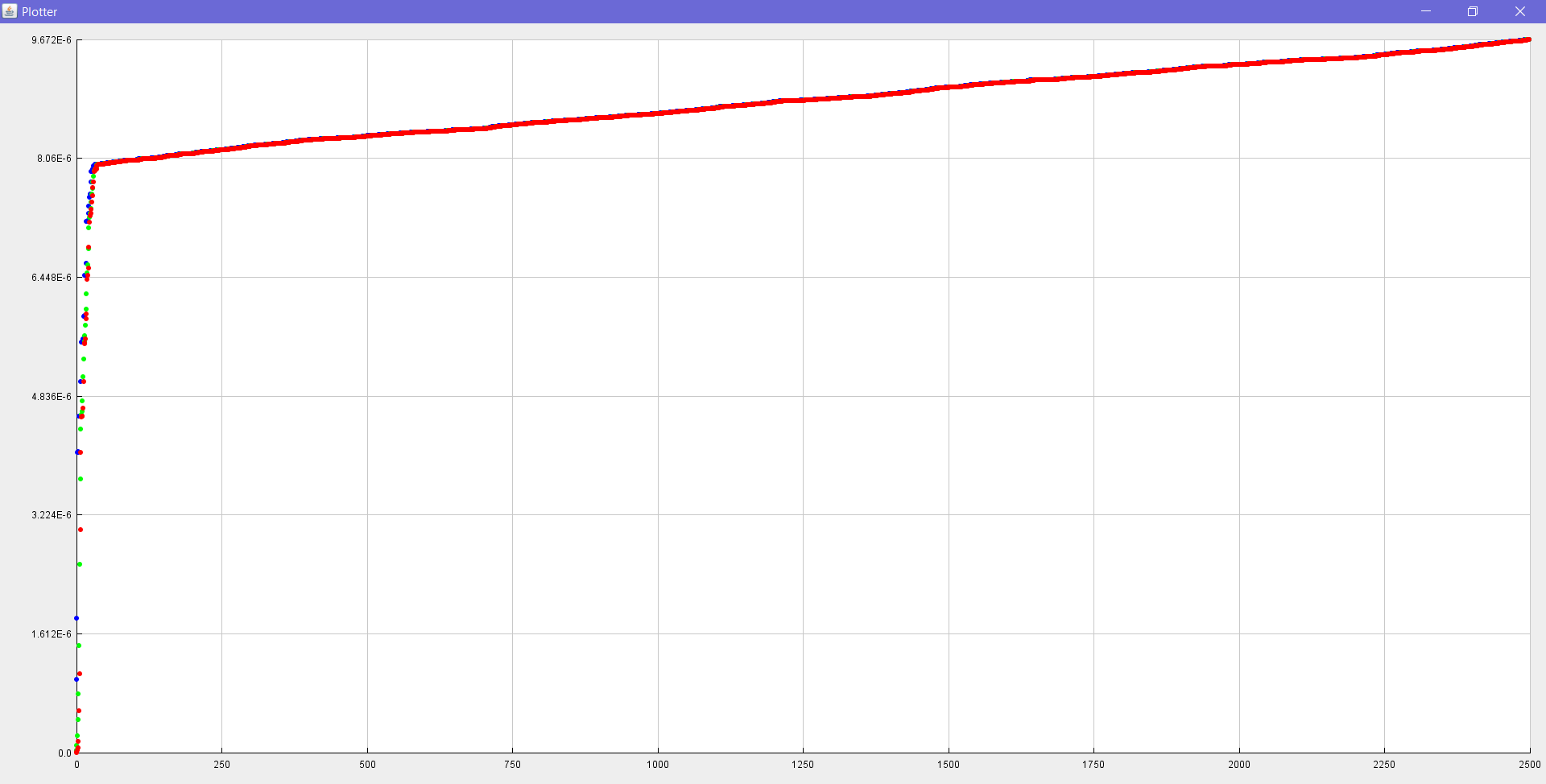
شکل 10 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 5

معادله‌ی پیشینی‌شده:

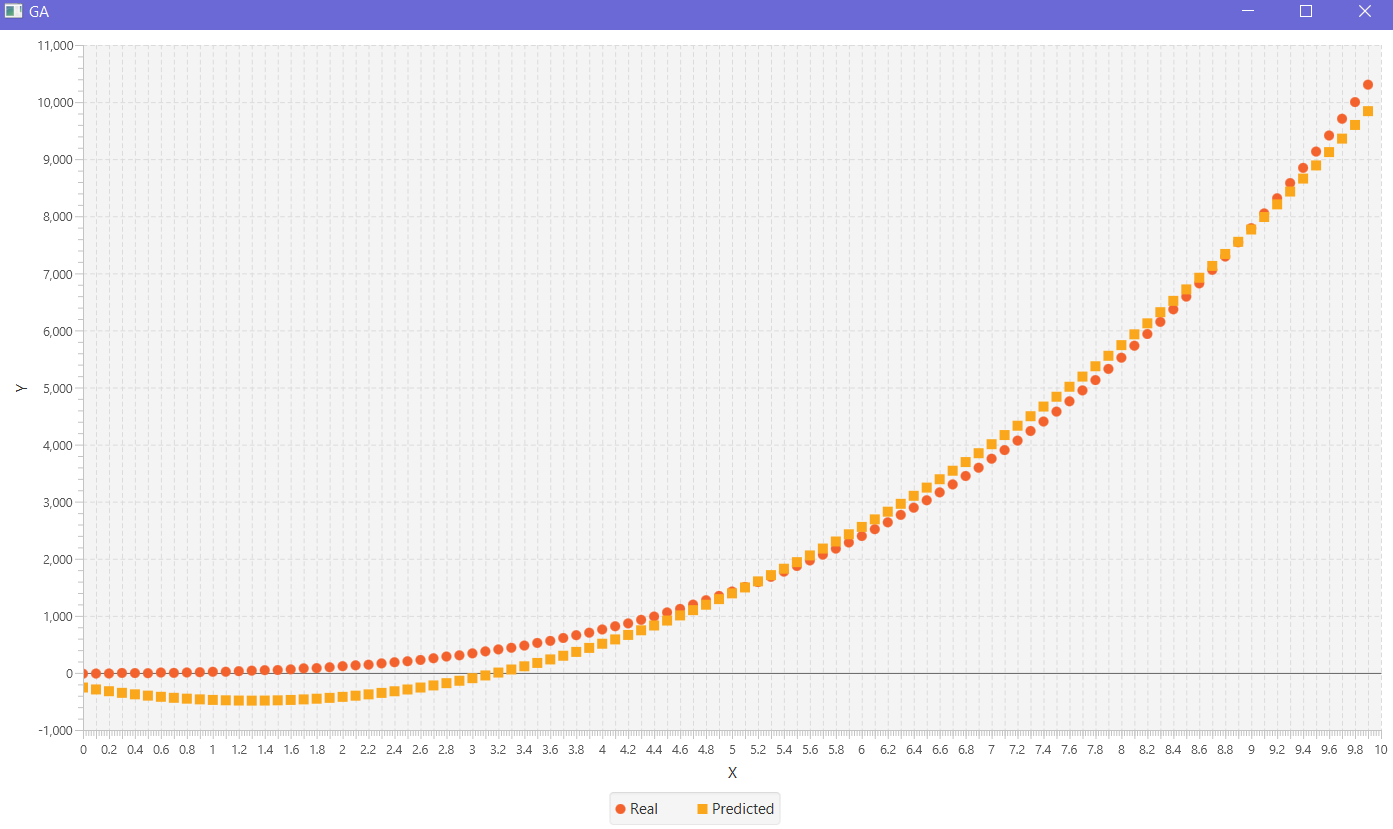
*تست (6) تأثیر اندازه‌ی جمعیت*

*NumberOfGenerations = 2500 , PopulationSize = 200, TournamentSize = 2, MutationRate = 0.02, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*

**

شکل 11 – نمودار بهترین، بدترین و میانگین شایستگی تست 6



شکل 12 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 6

معادله‌ی پیشینی‌شده:

**نتیجه‌گیری**

همانطور که از نتایج تست‌های (4)، (5) و (6) و جدول 2 مشخص است، می‌توان گفت اندازه‌ی جمعیت نسبت مستقیمی با ضرایب تولید شده نداشت.

|  |  |
| --- | --- |
| فیتنس آخرین نسل | اندازه‌ی جمعیت |
| 5.123E-6 | 50 |
| 2.9651E-5 | 100 |
| 9.672E-6 | 200 |

جدول 2 – تأثیر اندازه‌ی جمعیت بر شایستگی نسل آخر

احتمال انتخاب تصادفی یک individual با شایستگی خوب ربطی به اندازه‌ی جمعیت ندارد. در نتیجه نمی‌توان تأثیر مستقیمی بین این متقیر با میزان همگرایی رابطه‌ی مستقیمی پیدا کرد.

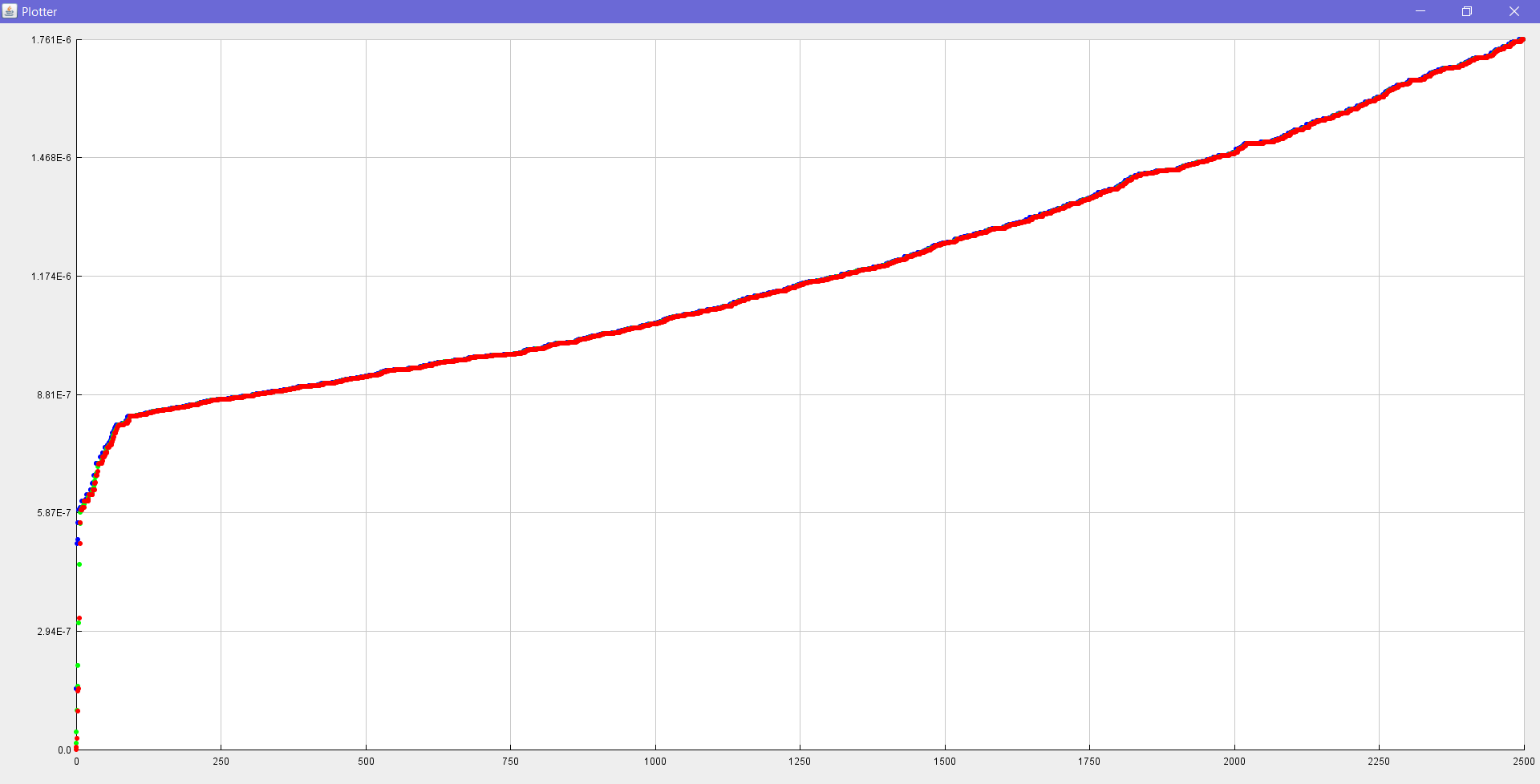
**تأثیر متغیر Tournament Size**

برای تشخیص تأثیر Tournament Size بر این مسئله، کافیست این متغیر را برابر مقادیر 2، 5 و 10 قرار دهیم و بقیه‌ی متغیرها را ثابت فرض کنیم. برای کاهش تاثیر جمعیت اولیه‌ی تصادفی، برای هر 3 مقدار از جمعیت اولیه‌ی تصادفی یکسانی استفاده می‌کنیم.

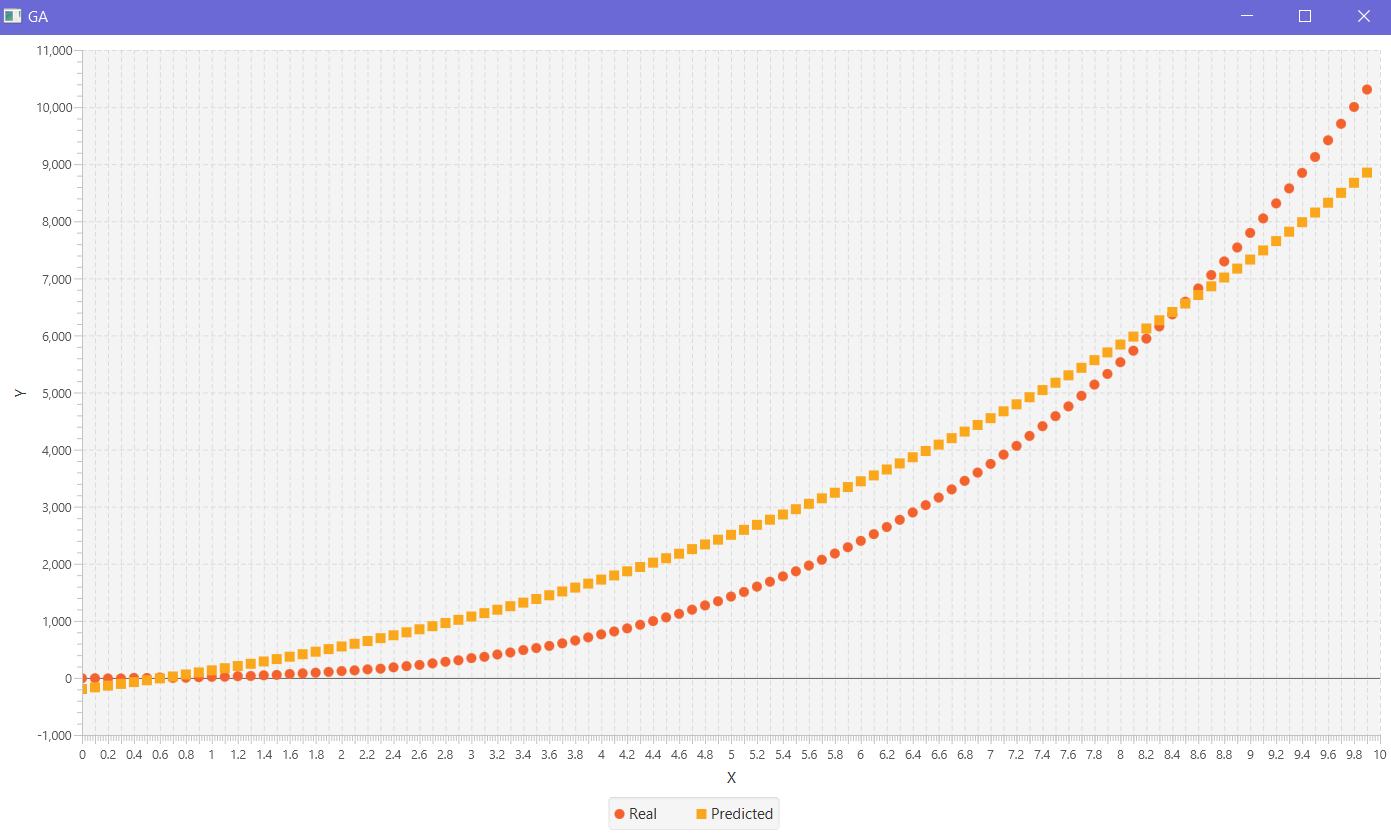
*تست (7) تأثیر Tournament Size*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.02, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*

**

شکل 13 – نمودار بهترین، بدترین و میانگین شایستگی تست 7



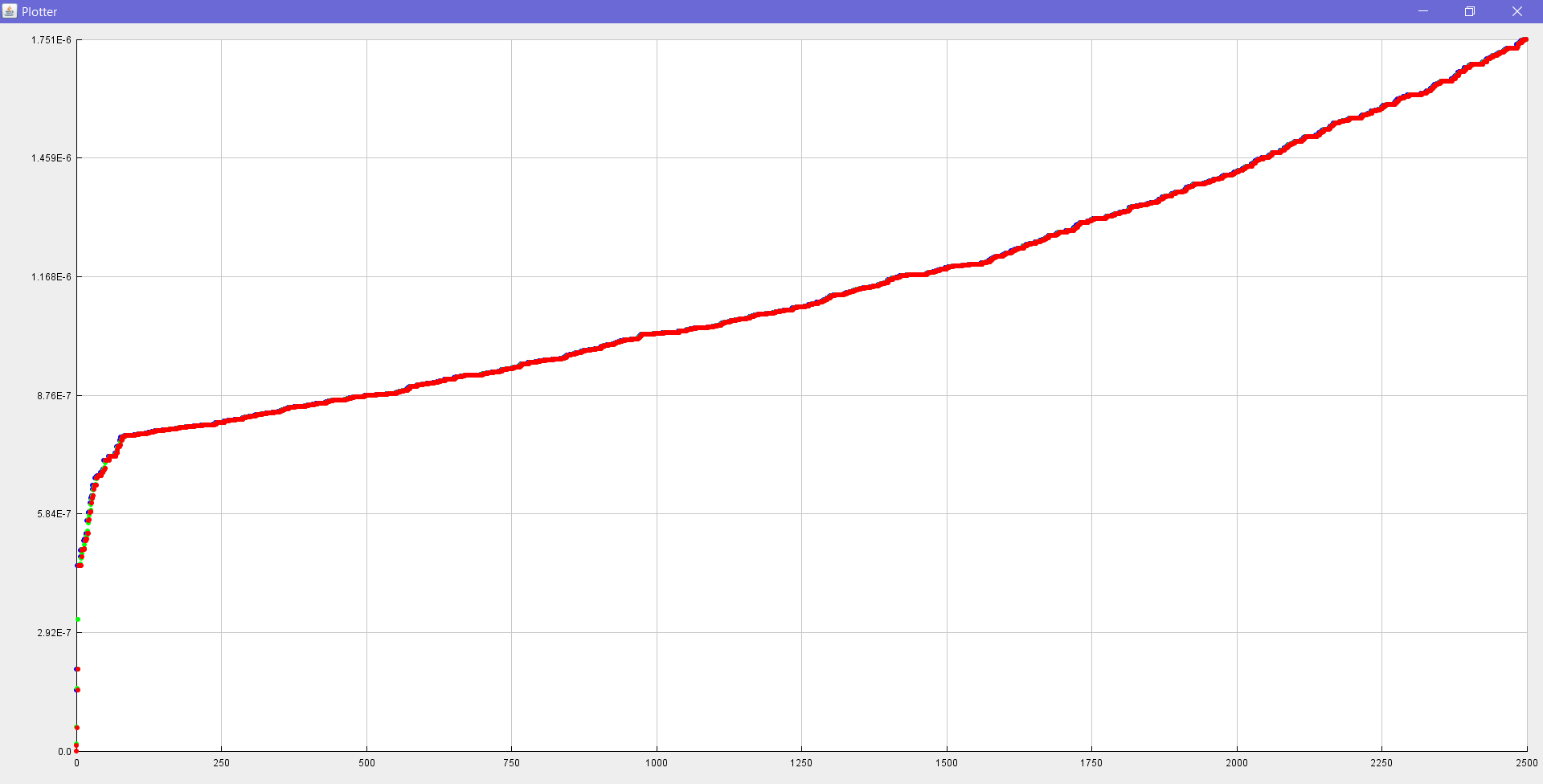
شکل 14 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 7

معادله‌ی پیشینی‌شده:

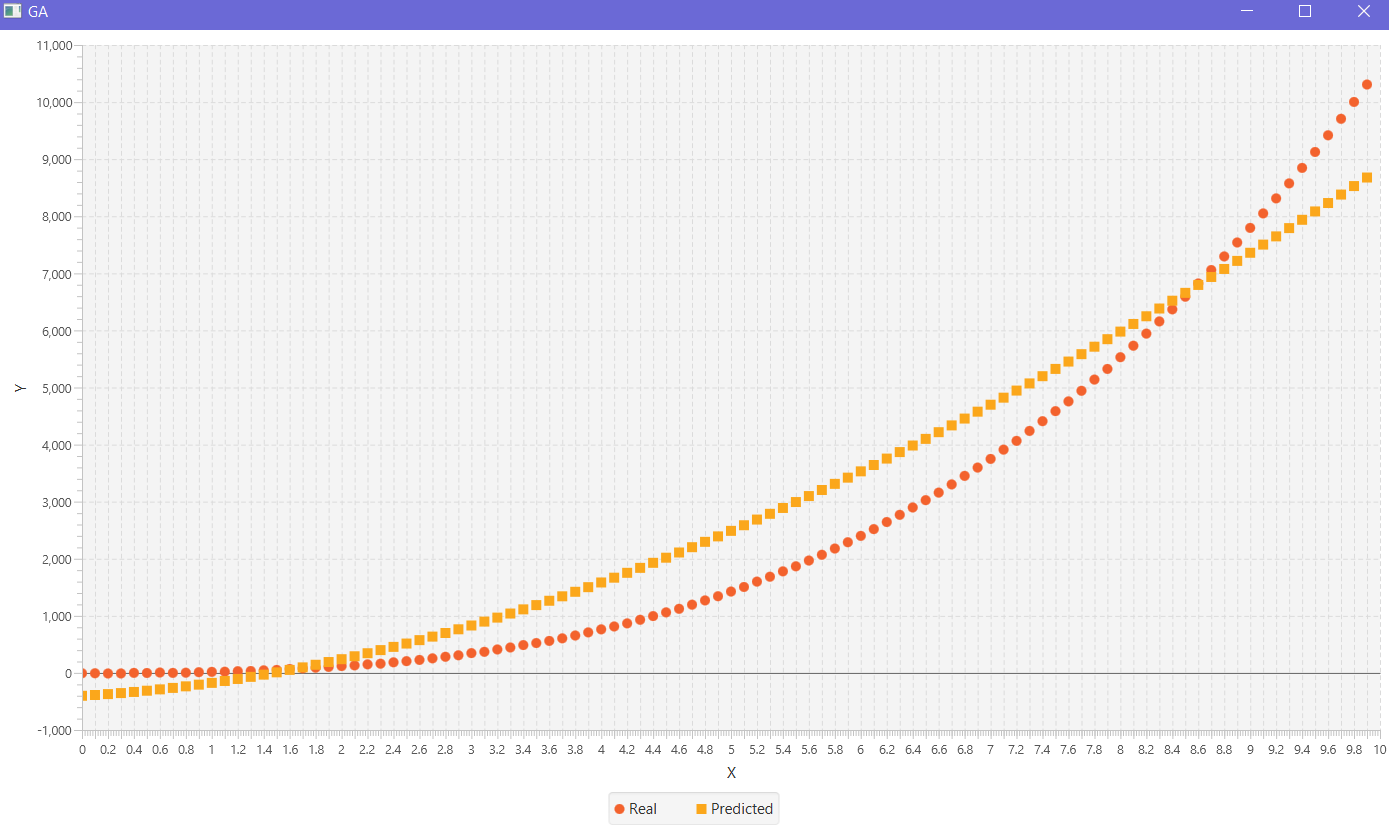
*تست (8) تأثیر Tournament Size*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 5, MutationRate = 0.02, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*



شکل 15 – نمودار بهترین، بدترین و میانگین شایستگی تست 8



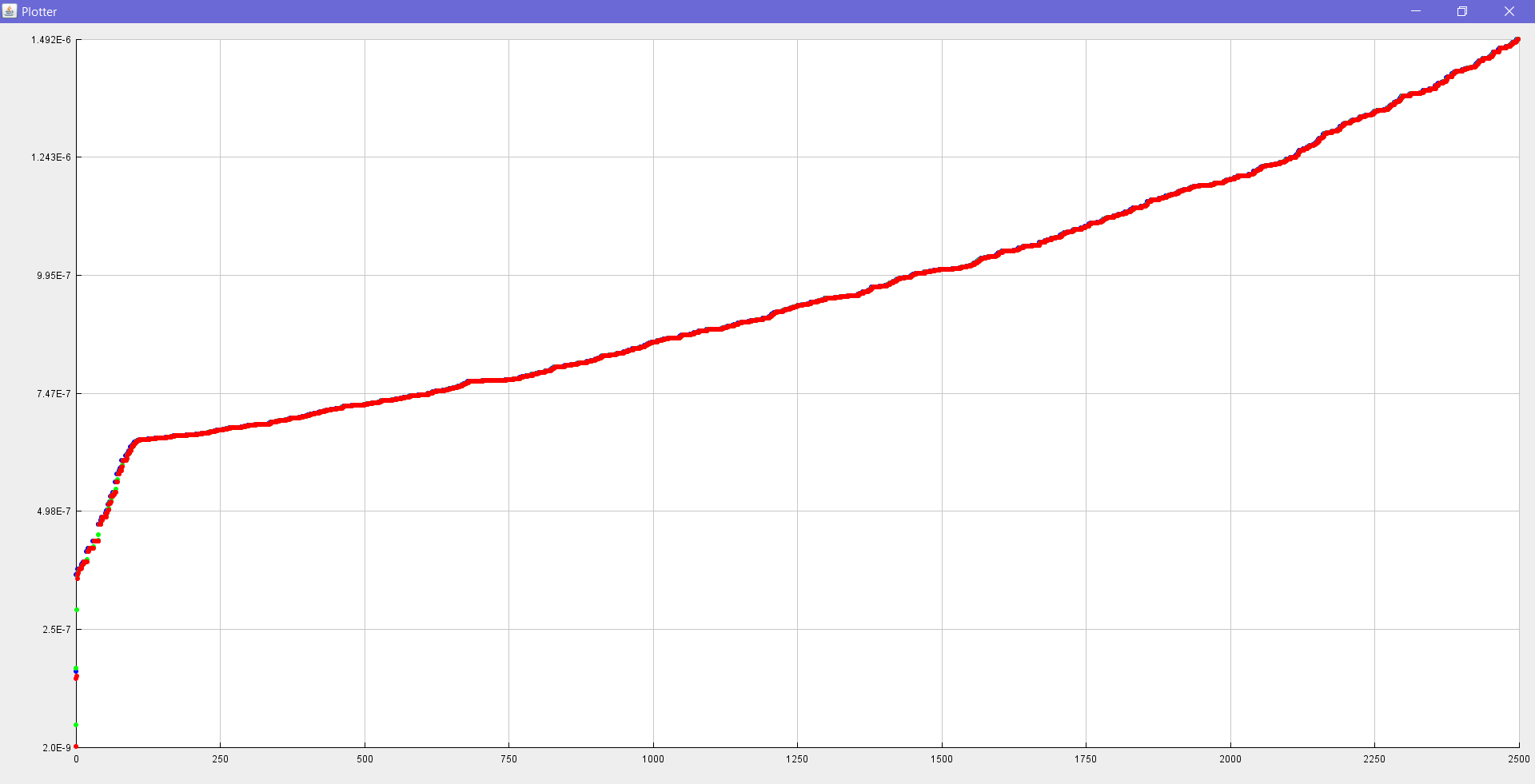
شکل 16 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 8

معادله‌ی پیشینی‌شده:

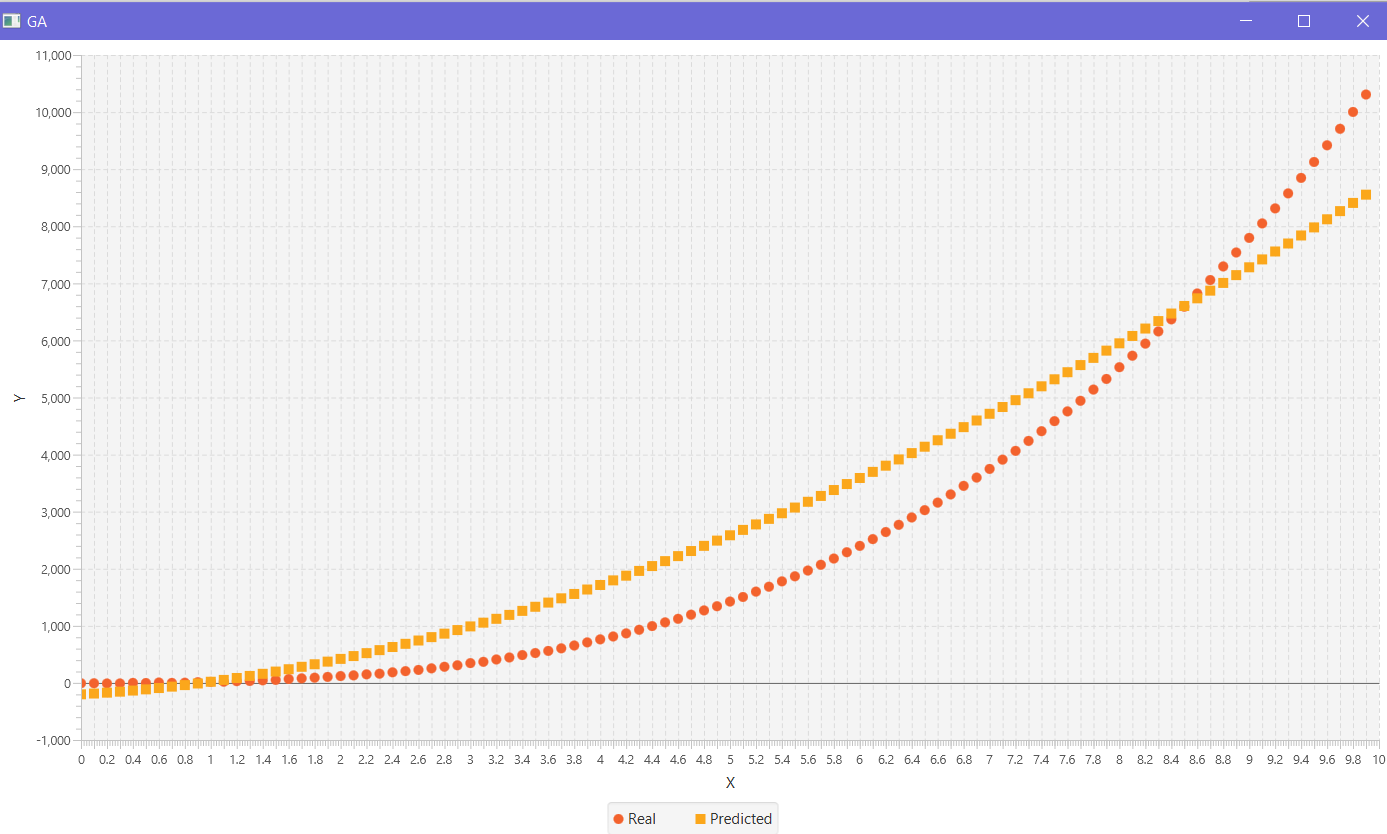
*تست (9) تأثیر Tournament Size*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 10, MutationRate = 0.02, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*

**

شکل 17 – نمودار بهترین، بدترین و میانگین شایستگی تست 9



شکل 18 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 9

معادله‌ی پیشینی‌شده:

**نتیجه‌گیری**

با توجه به تست‌های (7)، (8) و (9) و شکل‌های 13 و 15 و 17 می‌توان نتیجه گرفت با افزایش Tournament Size جمعیت سریعتر همگرا می‌شوند و این متغیر تأثیر مستقیمی بر جواب ما ندارد.

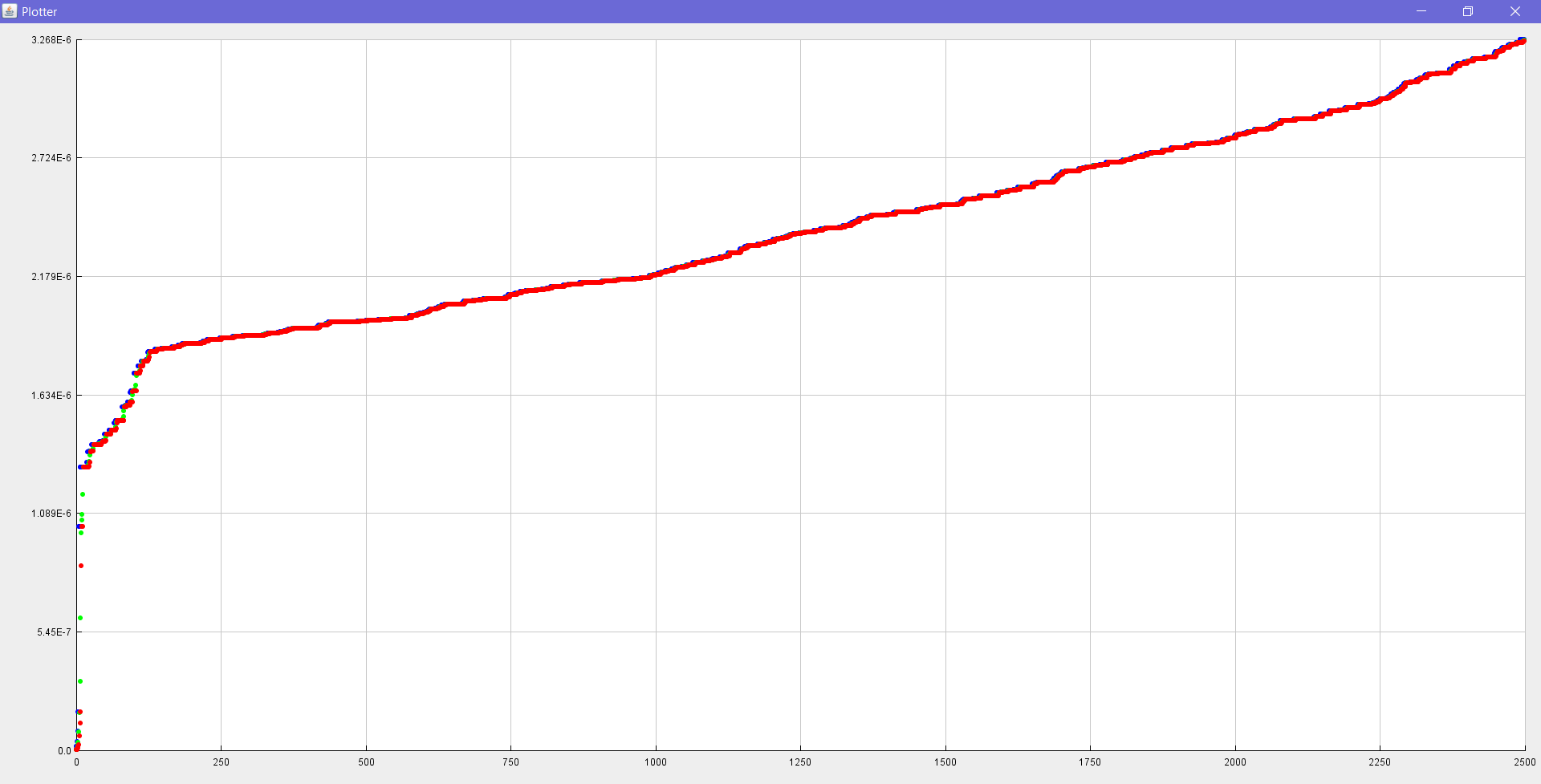
**تأثیر متغیر Mutation Rate**

برای تشخیص تأثیر متغیر Mutation Rate بر مسئله، این متغیر را برابر مقادیر 0.01، 0.02، 0.05 و 0.1 قرار دهیم و بقیه‌ی متغیرها را ثابت فرض کنیم. برای کاهش تاثیر جمعیت اولیه‌ی تصادفی، برای هر 4 مقدار از جمعیت اولیه‌ی تصادفی یکسانی استفاده می‌کنیم.

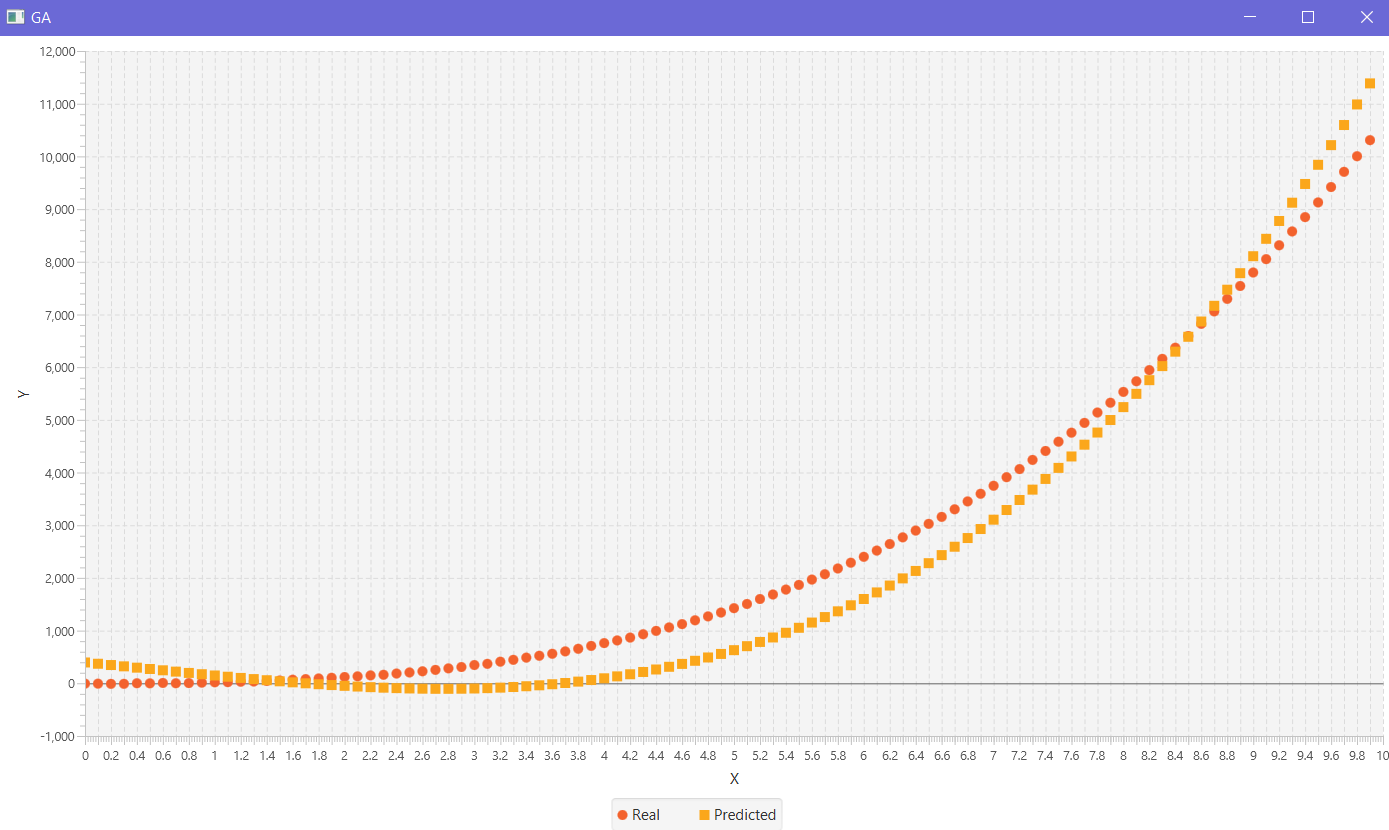
*تست (10) تأثیر Mutation Rate*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.01, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*

**

شکل 19 – نمودار بهترین، بدترین و میانگین شایستگی تست 10

**

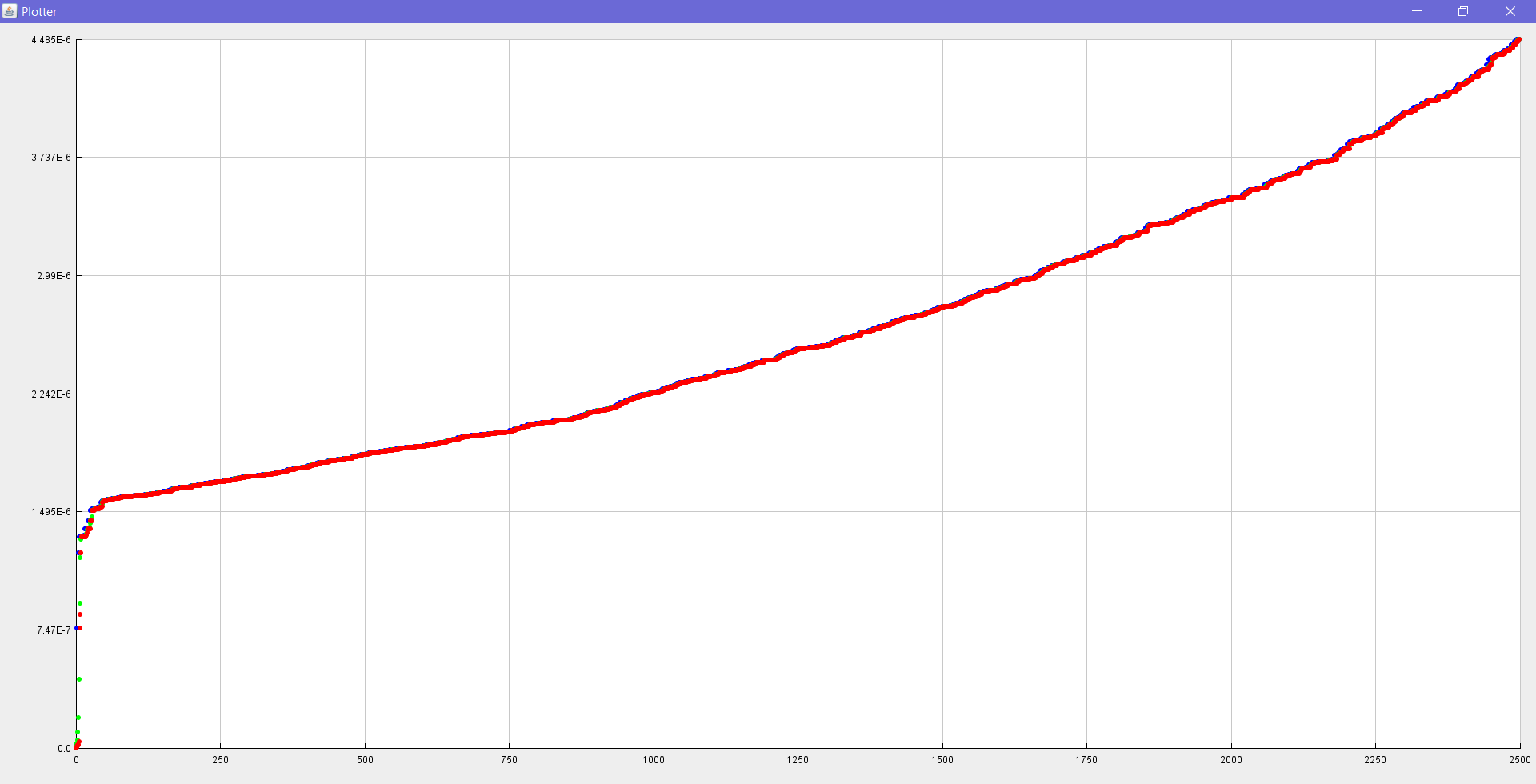
شکل 20 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 10

معادله‌ی پیشینی‌شده:

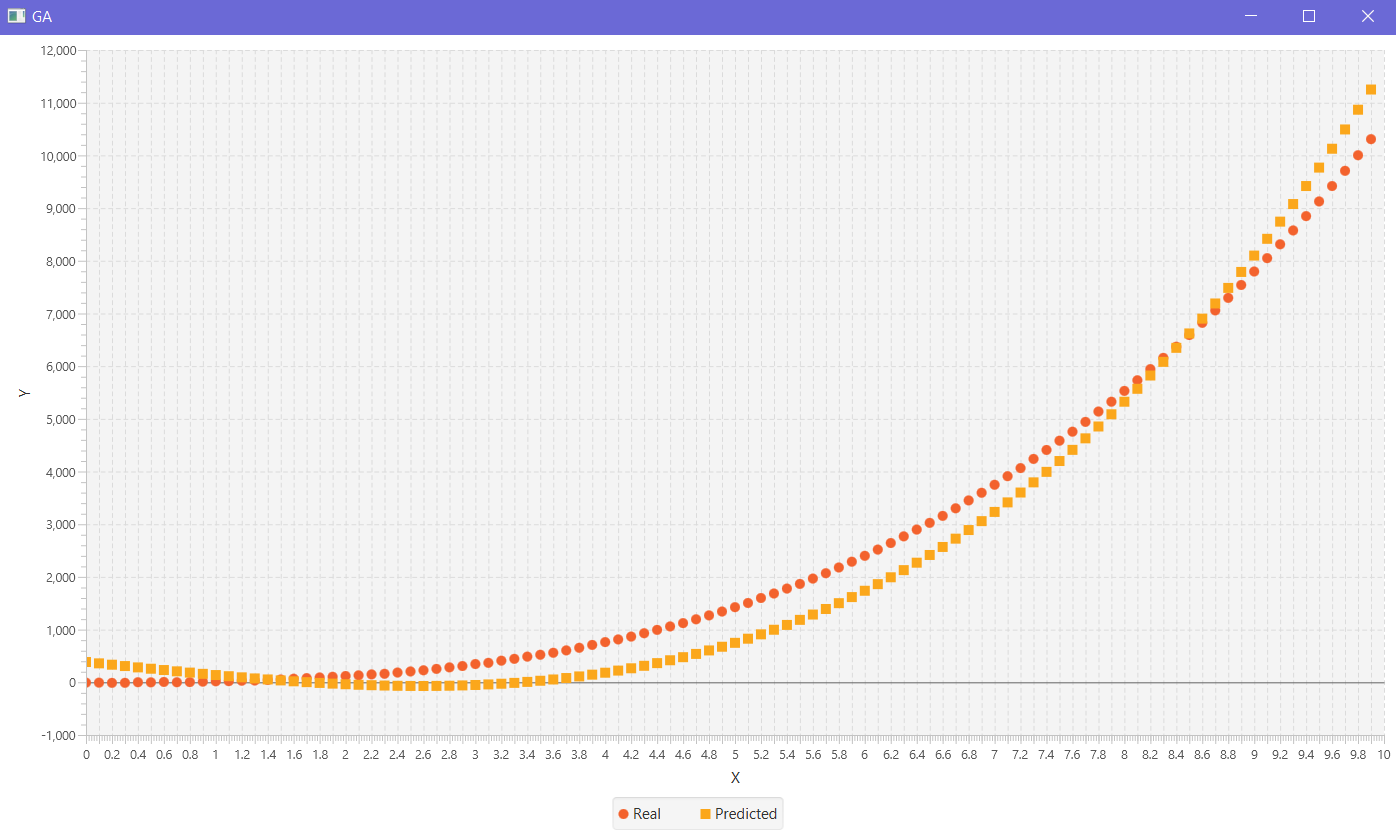
*تست (11) تأثیر Mutation Rate*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.02, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*

****

شکل 21 – نمودار بهترین، بدترین و میانگین شایستگی تست 11

****

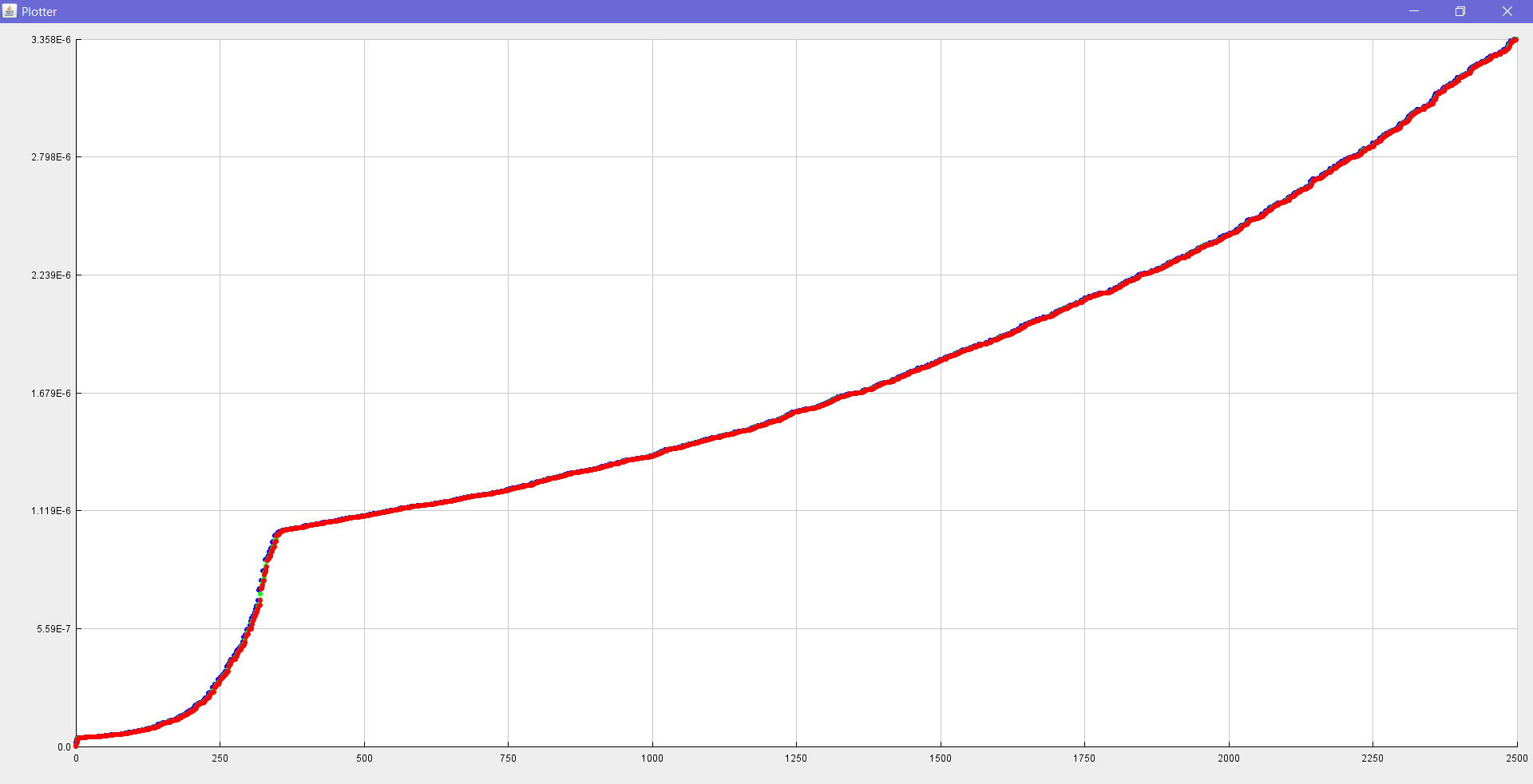
شکل 22 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 11

معادله‌ی پیشینی‌شده:

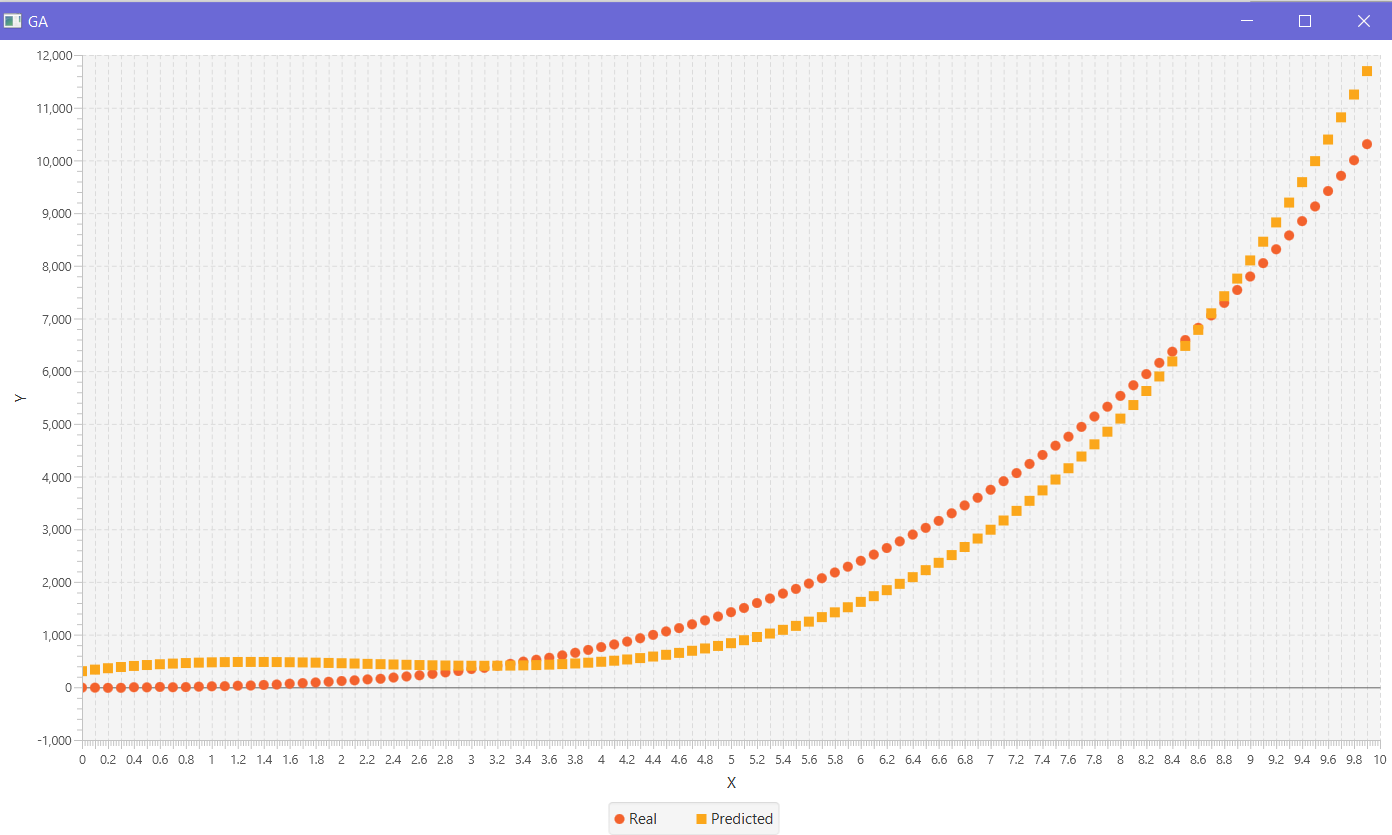
*تست (12) تأثیر Mutation Rate*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.05, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*

****

شکل 23 – نمودار بهترین، بدترین و میانگین شایستگی تست 12

****

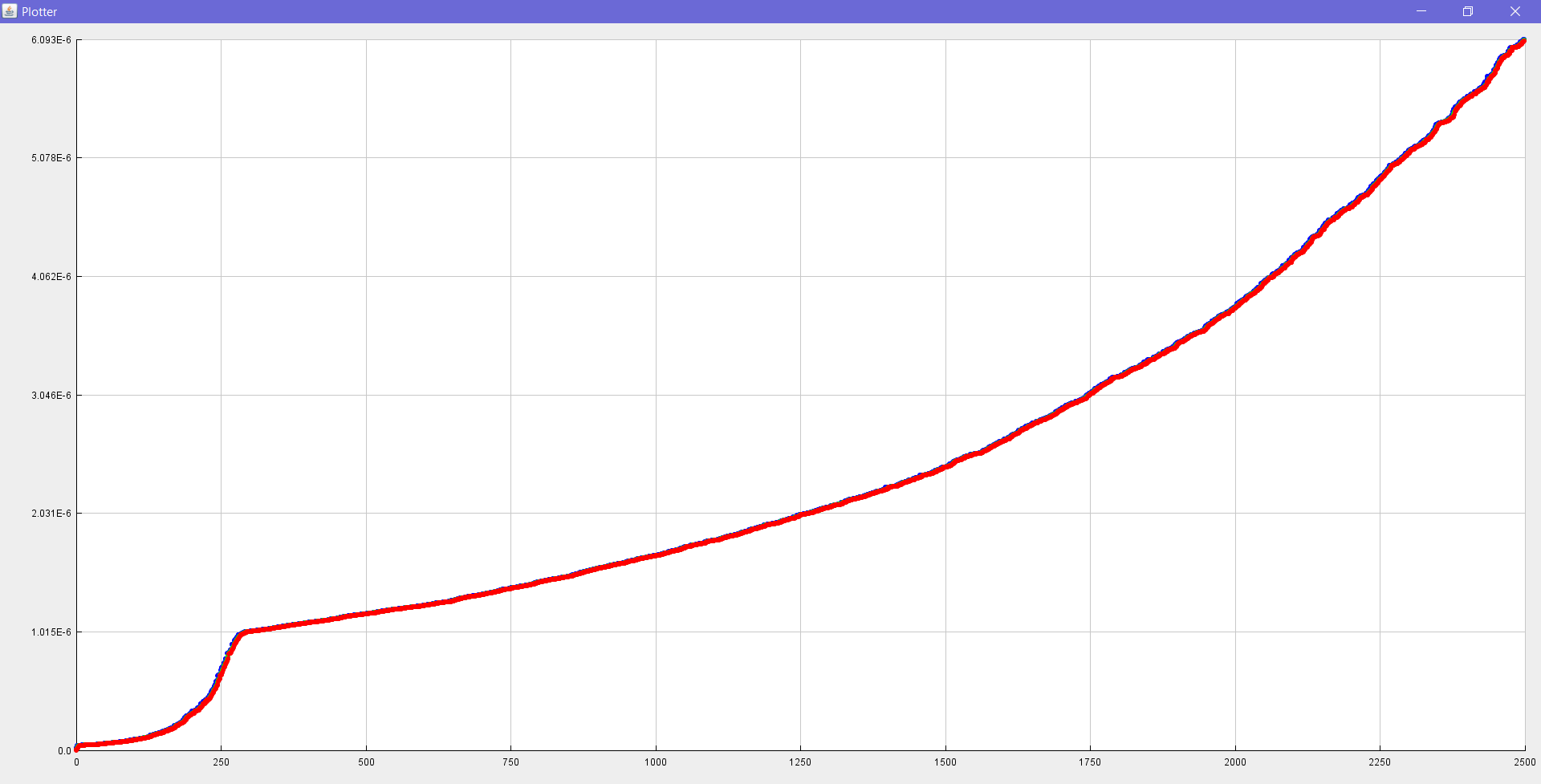
شکل 24 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 12

معادله‌ی پیشینی‌شده:

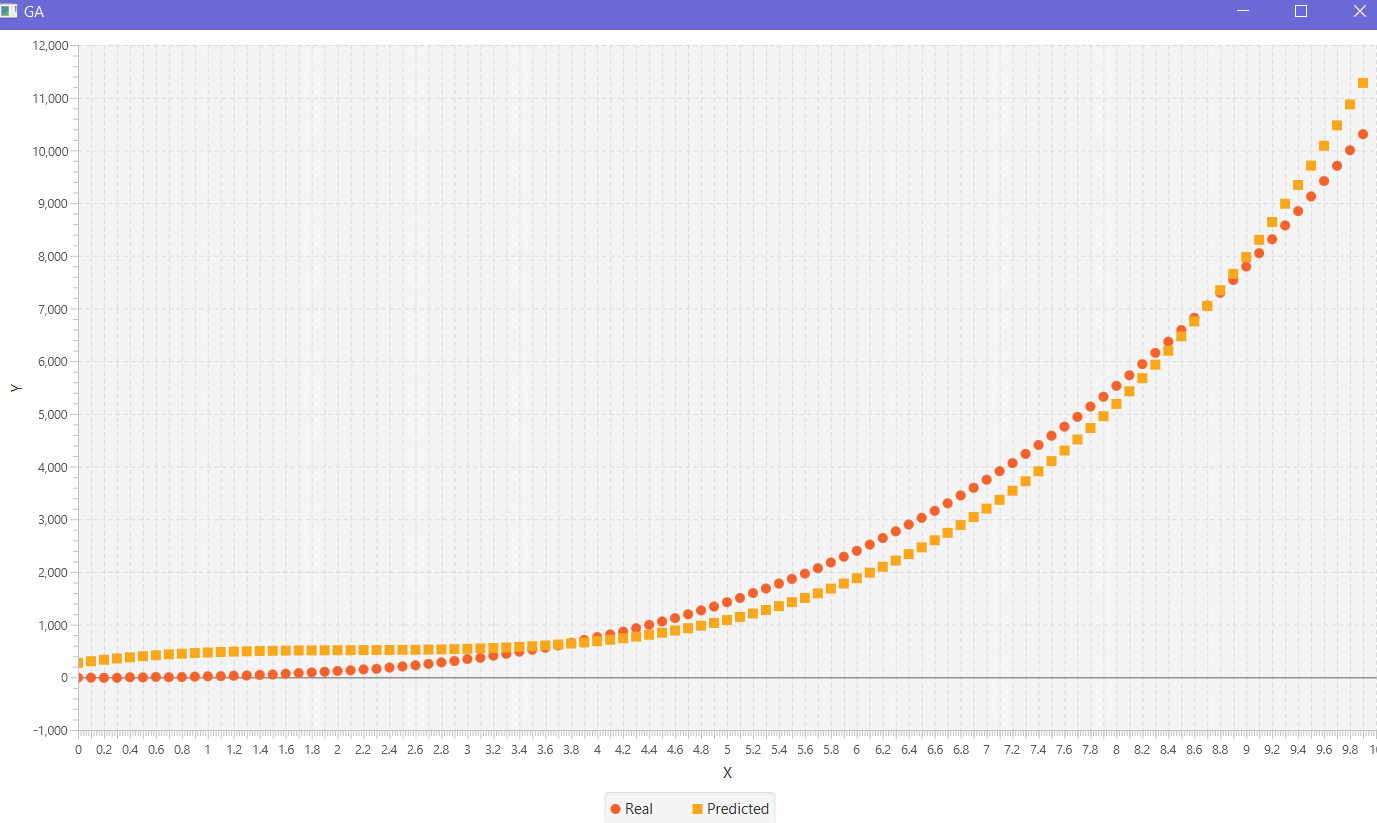
*تست (13) تأثیر Mutation Rate*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.1, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*

****

شکل 25 – نمودار بهترین، بدترین و میانگین شایستگی تست 13

****

شکل 26 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 13

معادله‌ی پیشینی‌شده:

**نتیجه‌گیری**

با توجه به تست‌های (10)، (11)، (12) و (13) می‌توان نتیجه گرفت که با افزایش نرخ جهش، جواب‌ها نیز بهتر می‌شوند. هر چه نرخ جهش بیشتر باشد، جهش هم بیشتر می‌شود. در نتیجه امکان اتفاق افتادن جهش مناسب هم بیشتر می‌شود. همچنین جهش‌هایی که موجب ضعیف شدن individual ها میشوند در مرحله‌ی بازگشت از بین می‌روند.

|  |  |
| --- | --- |
| فیتنس آخرین نسل | نرخ جهش |
| 3.268E-6 | 0.01 |
| 4.485E-6 | 0.02 |
| 3.358E-6 | 0.05 |
| 6.093E-6 | 0.1 |

جدول 3

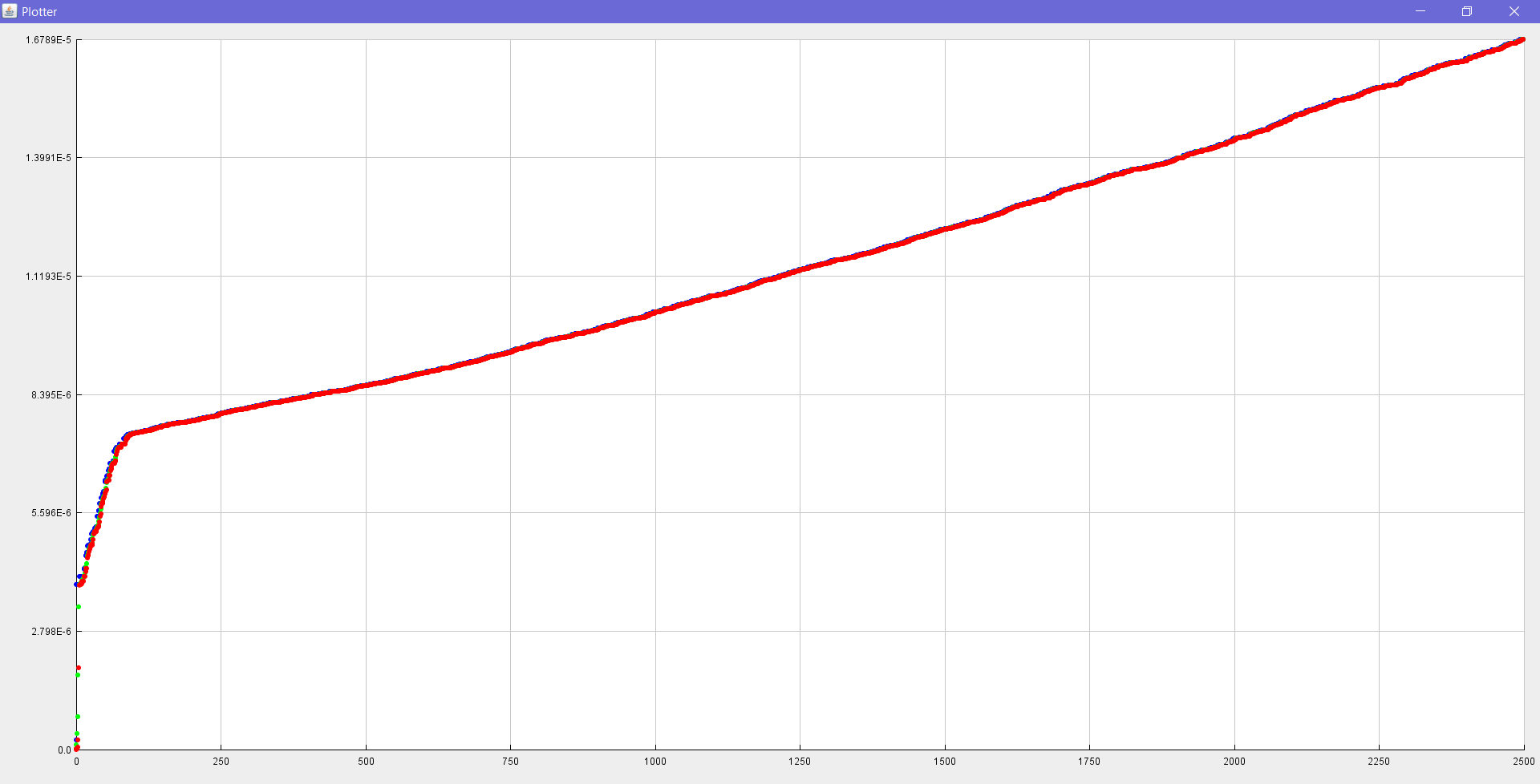
**تاثیر متغیر Variance**

برای تشخیص تأثیر متغیر Variance بر مسئله، این متغیر را برابر مقادیر 0.001، 0.01، 0.1، 1 و 10 قرار می‌دهیم و بقیه‌ی متغیرها را ثابت فرض می‌کنیم. برای کاهش تاثیر جمعیت اولیه‌ی تصادفی، برای هر 4 مقدار از جمعیت اولیه‌ی تصادفی یکسانی استفاده می‌کنیم.

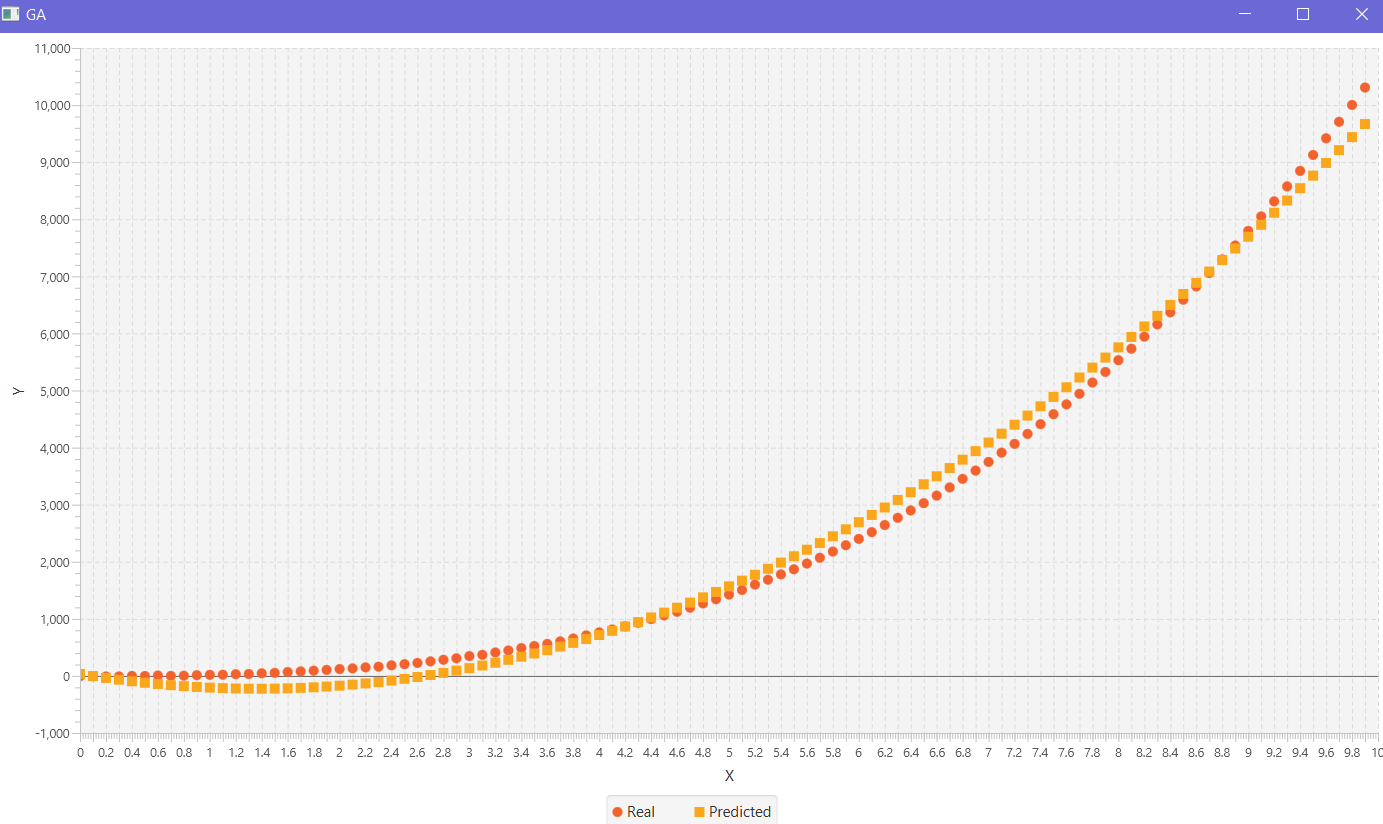
*تست (14) تأثیر Variance*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.05, Variance = 0.001*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*



شکل 27 – نمودار بهترین، بدترین و میانگین شایستگی تست 14



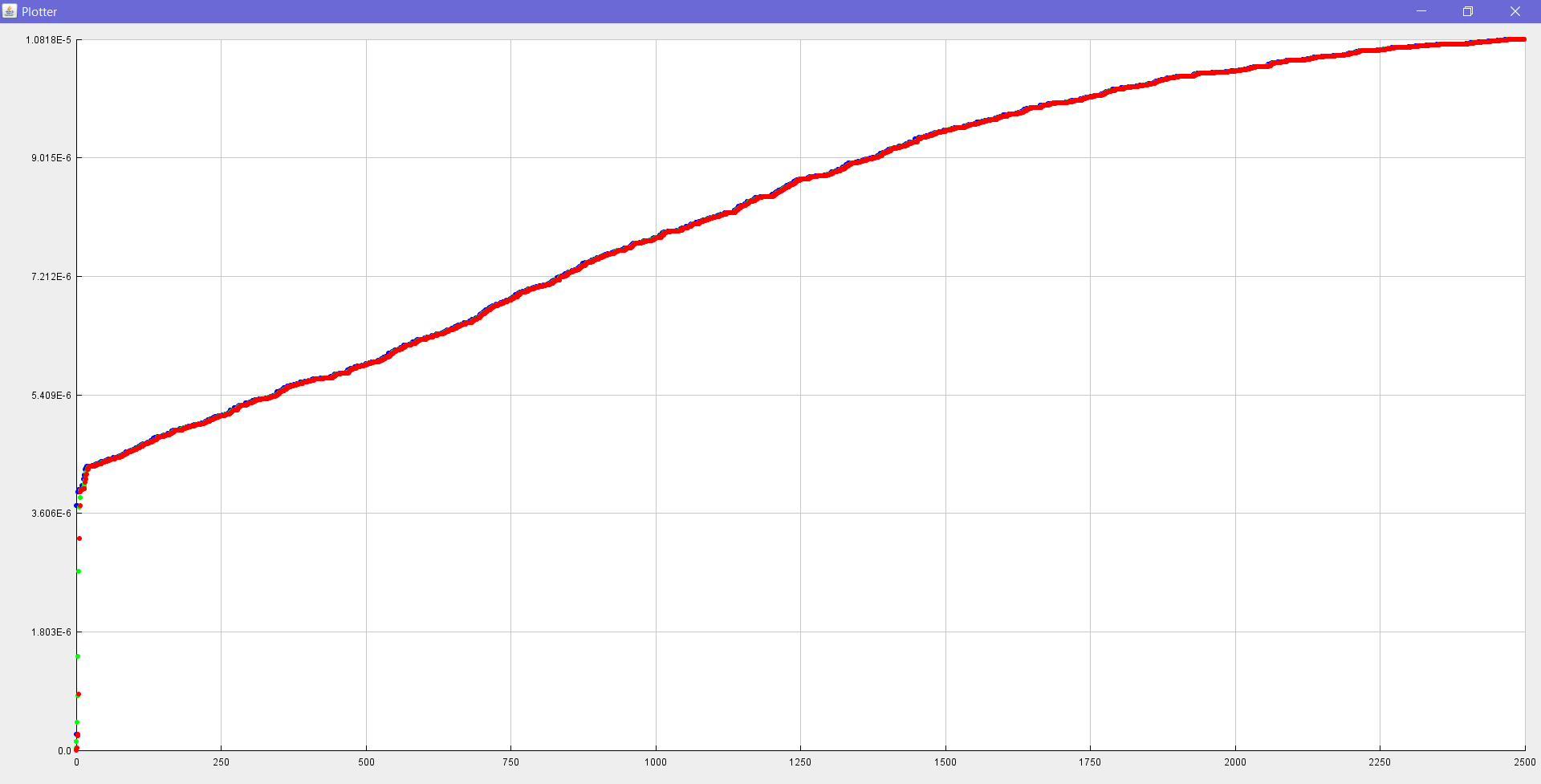
شکل 28 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 14

معادله‌ی پیشینی‌شده:

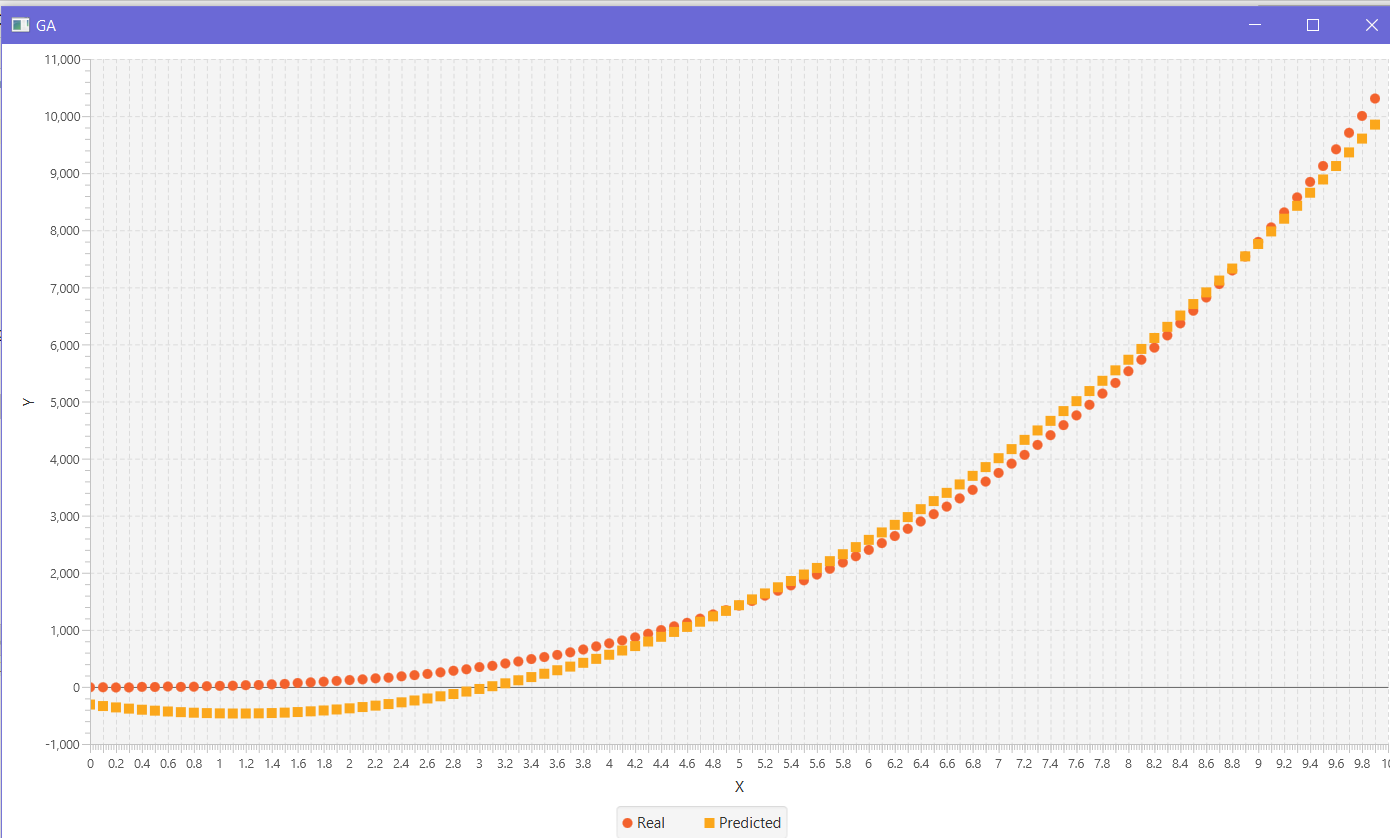
*تست (15) تأثیر Variance*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.05, Variance = 0.01*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*



شکل 29 – نمودار بهترین، بدترین و میانگین شایستگی تست 15



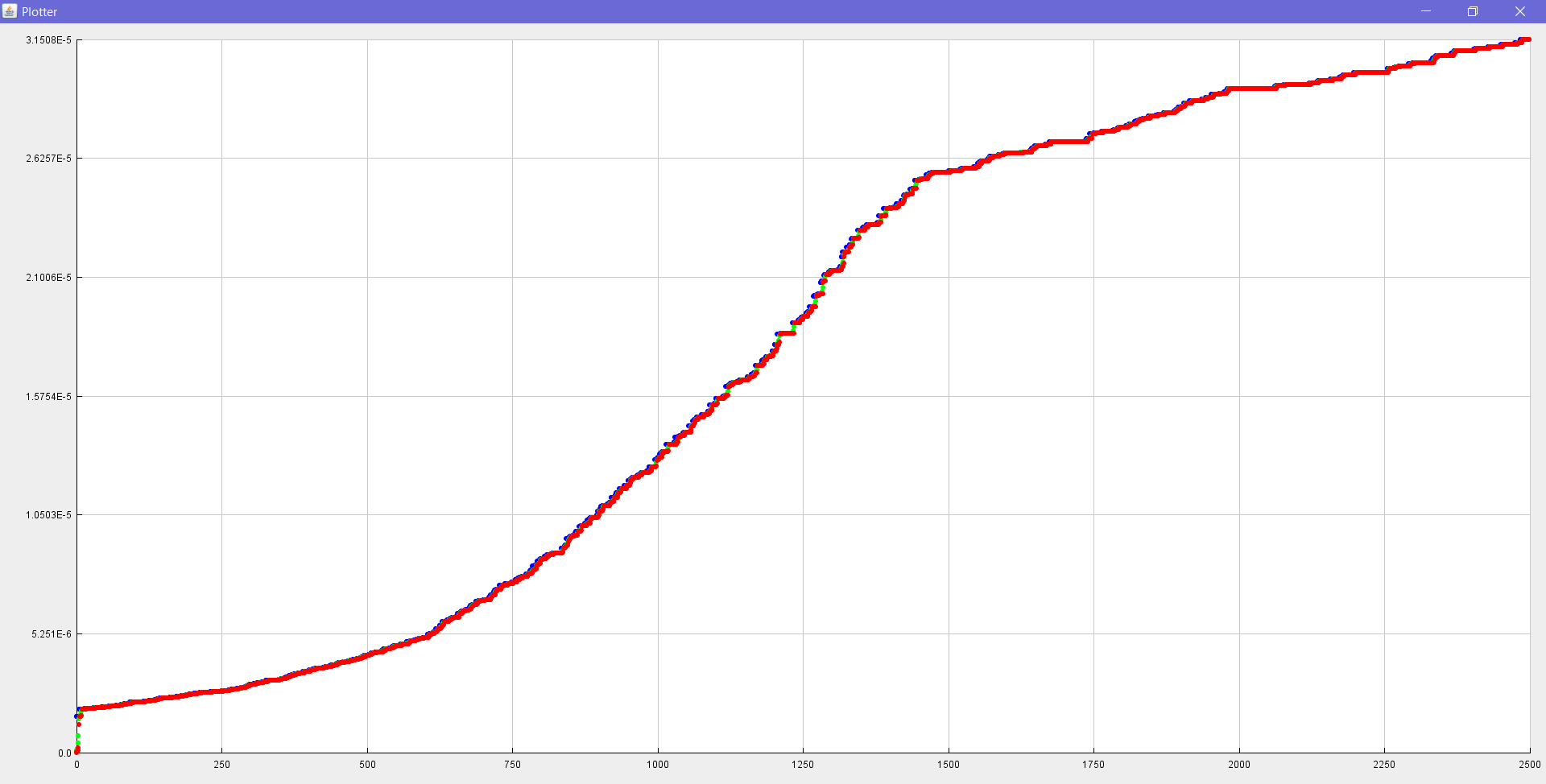
شکل 30 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 15

معادله‌ی پیشینی‌شده:

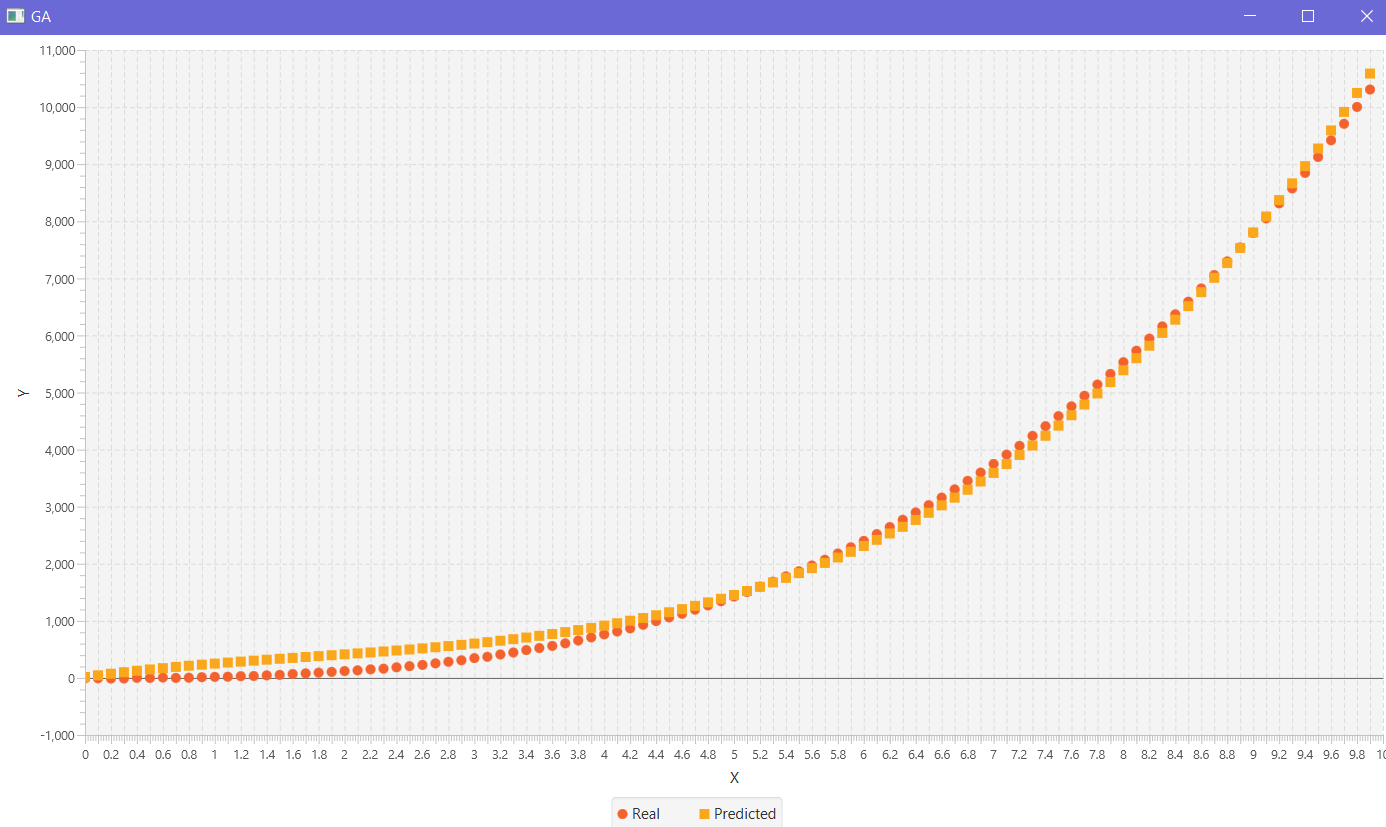
*تست (16) تأثیر Variance*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.05, Variance = 0.1*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*



شکل 31 – نمودار بهترین، بدترین و میانگین شایستگی تست 16



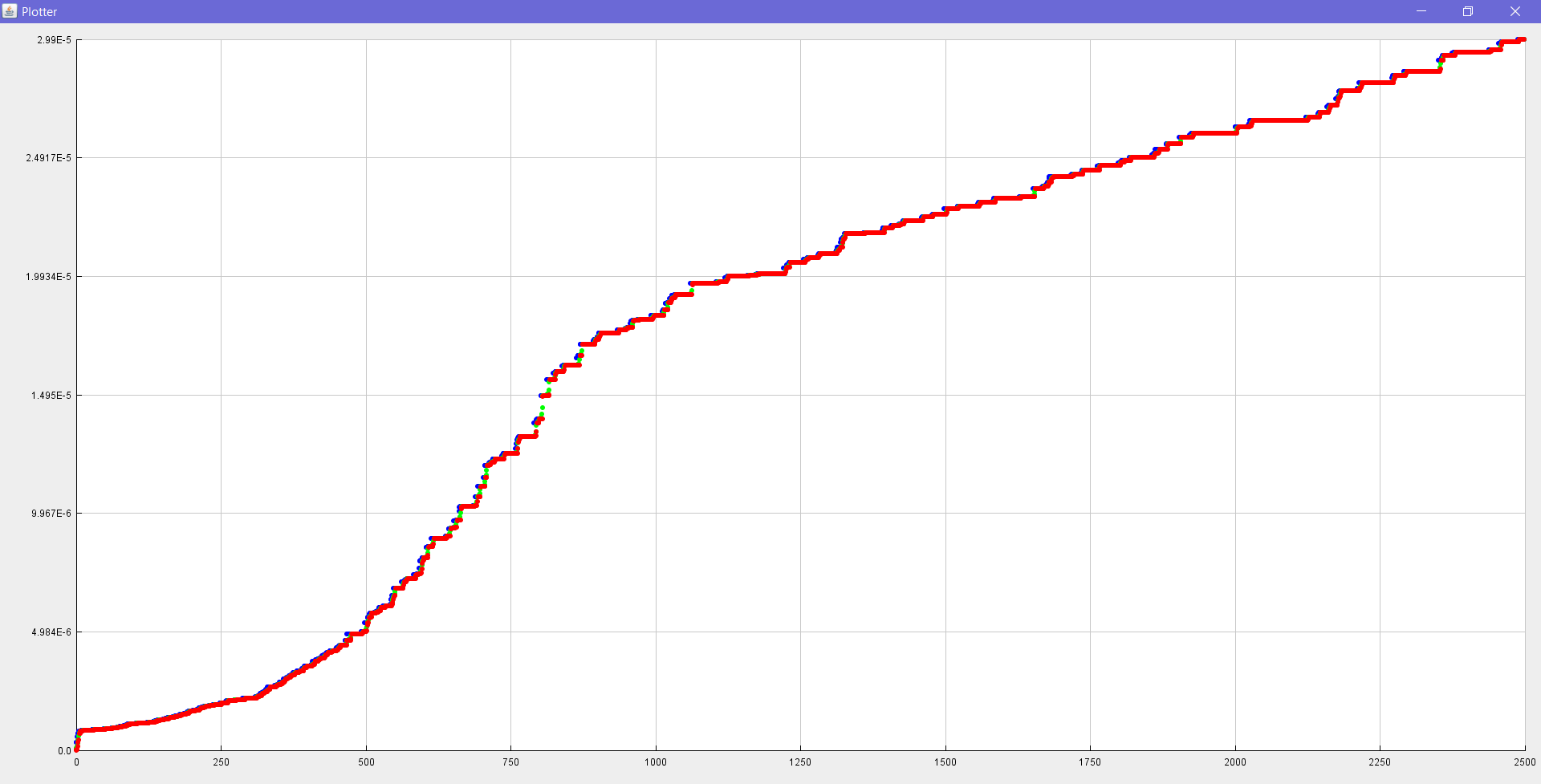
شکل 32 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 16

معادله‌ی پیشینی‌شده:

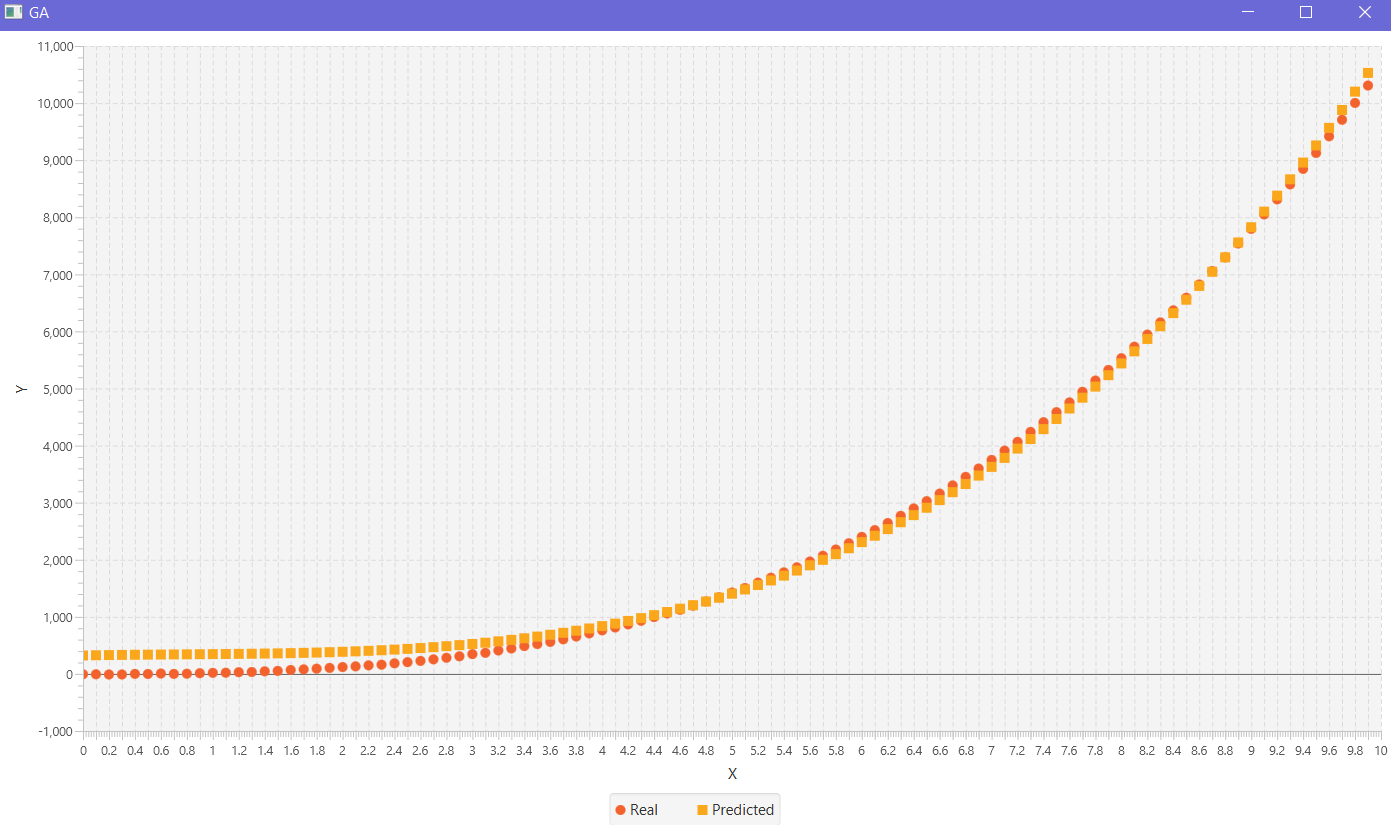
*تست (17) تأثیر Variance*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.05, Variance = 1*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*



شکل 33 – نمودار بهترین، بدترین و میانگین شایستگی تست 17



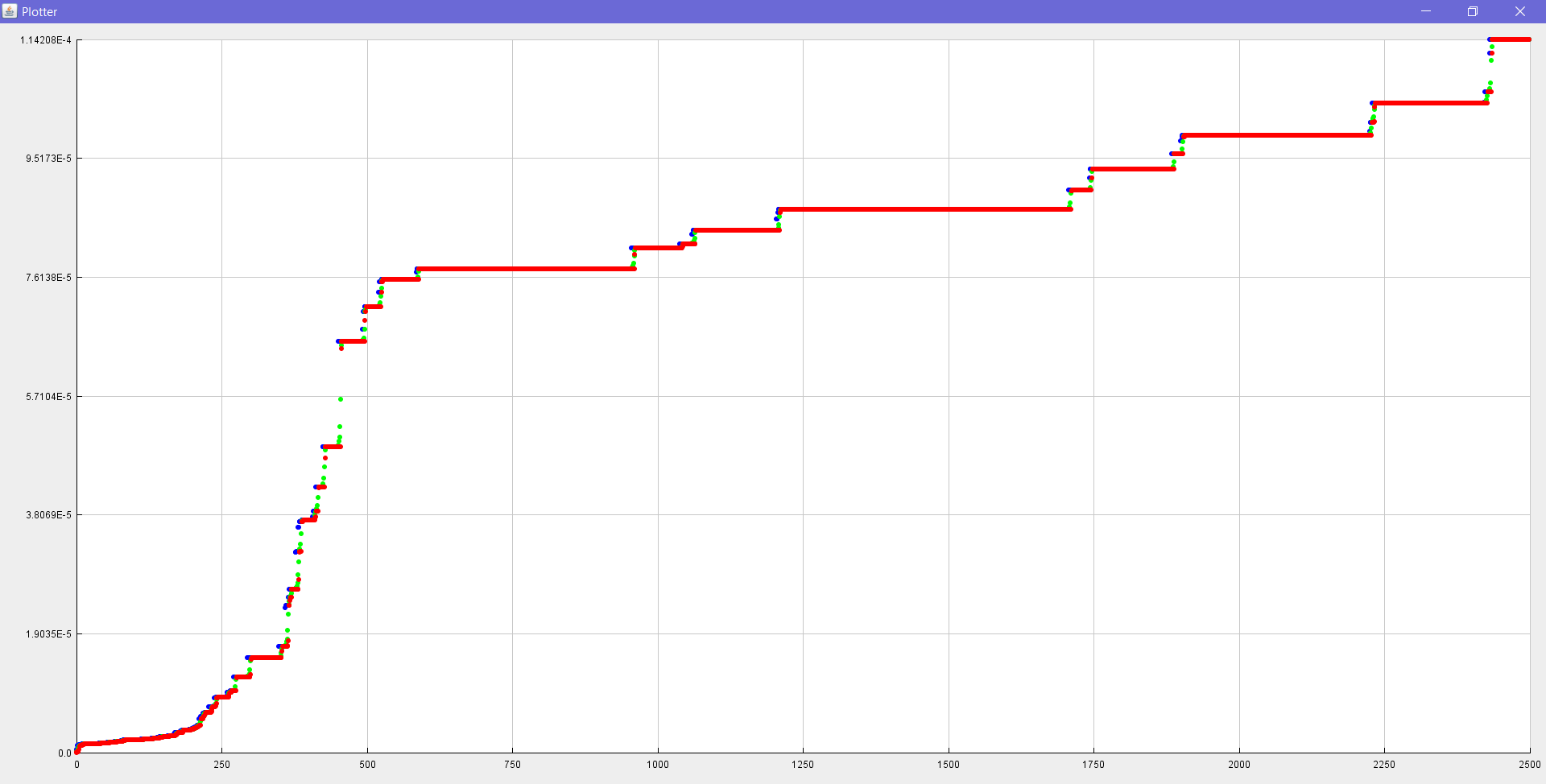
شکل 34 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 17

معادله‌ی پیشینی‌شده:

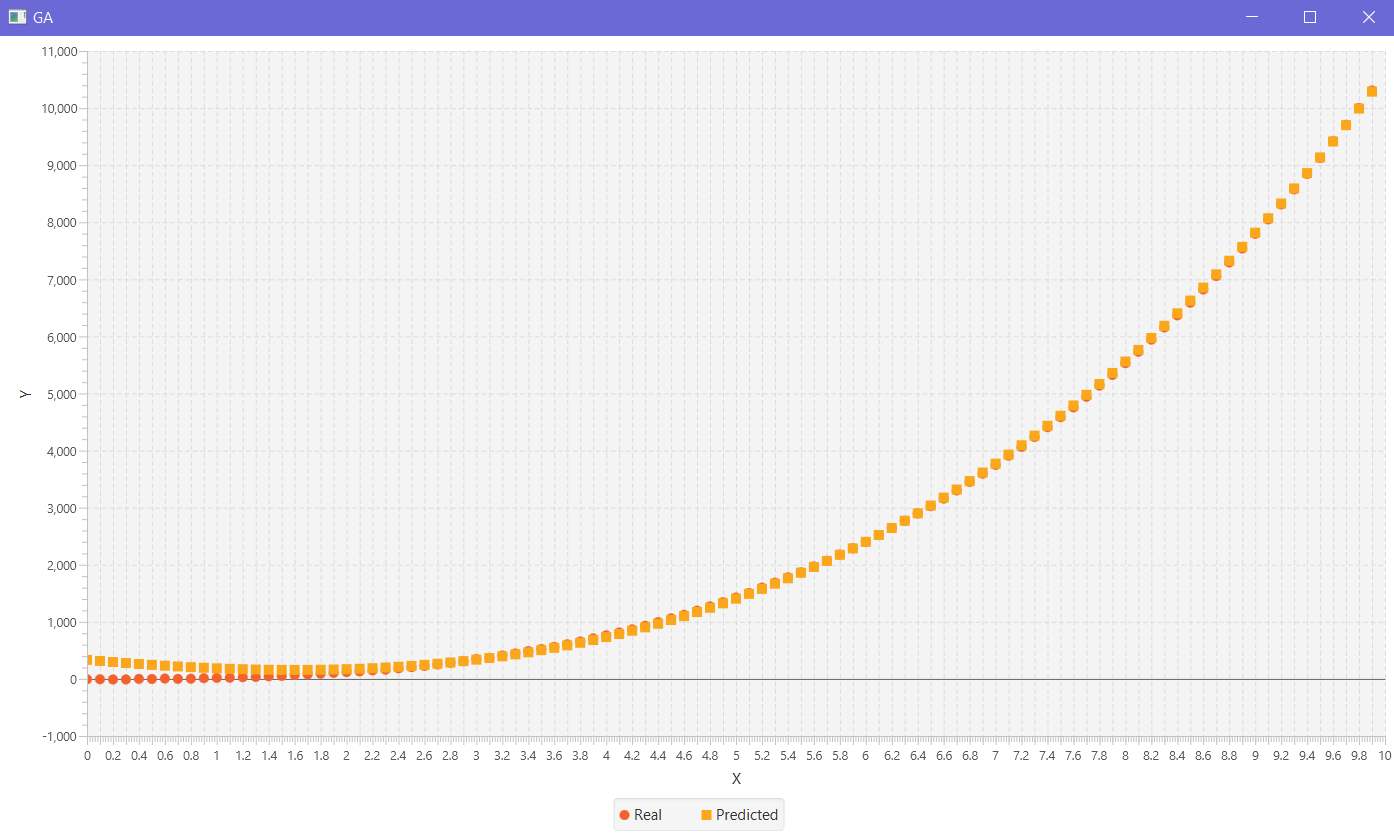
*تست (18) تأثیر Variance*

*NumberOfGenerations = 2500 , PopulationSize = 50, TournamentSize = 2, MutationRate = 0.05, Variance = 10*

*در این تست بازه‌ی انتخاب تصادفی ضرایب بین -400 تا 400 است.*



شکل 35 – نمودار بهترین، بدترین و میانگین شایستگی تست 18



شکل 36 – مقایسه‌ی حالت پیشبینی شده با واقعیت تست 18

معادله‌ی پیشینی‌شده:

**نتیجه‌گیری**

با توجه به نتایج تست‌های (14)، (15)، (16)، (17) و (18) و جدول 4 می‌توان نتیجه گرفت که با افزایش Variance نتایج بهتری می‌توان کسب کرد.

|  |  |
| --- | --- |
| فیتنس آخرین نسل | Variance |
| 1.6789E-5 | 0.001 |
| 1.0818E-5 | 0.01 |
| 3.1508E-5 | 0.1 |
| 2.99E-5 | 1 |
| 1.14208E-4 | 10 |

جدول 4

همچنین با توجه به شکل‌های 27، 29، 31، 33 و 35 می‌توان فهمید که با افزایش Variance مقدار نویز گاوسی زیاد می‌شود. در نتیجه جهش‌ها با مقدار بیشتری انجام می‌شوند. به‌طوری که در شکل 35 و در شرایطی که Variance 10 لست نمودار حالتی پله‌گونه به خود گرفته است.