

НЕЙРОСЕТЕВАЯ МОДЕЛЬ УПРАВЛЯЮЩИХ ДЕЙСТВИЙ ЛЕТЧИКА¹

В задачах ручного управления применительно к замкнутой системе «самолет-летчик» важную роль играют математические модели характеристик управляющих действий летчика. Традиционные модели такого рода обладают целым рядом недостатков, которые затрудняют решение прикладных задач указанного класса. Для преодоления этих недостатков в работе построена нейросетевая модель (НС-модель) характеристик управляющих действий летчика. Дается способ построения таких моделей с использованием результатов экспериментальных исследований управляющих действий летчика.

Введение

Для решения многих прикладных задач ручного управления в замкнутой системе «самолет-летчик» [1] необходимо располагать математическими моделями характеристик управляющих действий летчика. Модели, обычно используемые для этих целей (модель частоты среза, оптимальная модель, структурная модель), имеют необходимую степень адекватности результатов моделирования лишь в ограниченном диапазоне частот [2-4], что ограничивает практическую ценность получаемых результатов. Кроме того, в упомянутых традиционных моделях отсутствуют возможности их адаптации к меняющимся условиям, в силу чего они могут быть использованы лишь для стационарных условий. Один из возможных путей преодоления отмеченных недостатков состоит в построении модели характеристик управляющих действий летчика, основанной на применении искусственных нейронных сетей (НС). Базой для построения такого рода НС-моделей служат результаты экспериментальных исследований управляющих действий летчика, полученные с помощью рабочей станции, специально предназначенной для изучения характеристик систем «самолет-летчик».

¹ Вестник МАИ. – 2007. том 14, №2. – с.53-66.

1. Формирование показателей, требуемых для построения нейросетевой модели управляющих действий летчика

Решение прикладных задач ручного управления базируется на использовании методов экспериментальных исследований, а также на математическом моделировании. Для оценки поведения летчика и формирования математических моделей характеристик его действий будем рассматривать замкнутую систему «самолет-летчик», показанную на рис.1 применительно к задаче одноконтурного компенсаторного слежения.



Рис.1. Замкнутая система «самолёт-лётчик» применительно к задаче компенсаторного слежения. Здесь $i(t)$ – входной сигнал, $y(t)$ – выходной сигнал, $c(t)$ – управляющий сигнал, $e(t)$ – сигнал ошибки (рассогласования)

Когда лётчик управляет объектом, выполняя данную задачу, он с помощью дисплея получает визуальную информацию о величине сигнала ошибки $e(t)$ и скорости его изменения. Перемещая рычаг управления и получая кинестетическую информацию, лётчик может оценить изменение выходного сигнала $y(t)$ и скорость этого изменения, то есть получить представление о реакции объекта управления на управляющий сигнал $c(t)$. В дальнейшем считается, что сигналы $e(t)$ и $c(t)$ – векторы, составляющими которых помимо значений собственно сигналов являются также и значения производных для рассматриваемых сигналов. Заменяя летчика его нейросетевой моделью в замкнутой системе «самолет-летчик», получим схему, изображенную на рис.2.

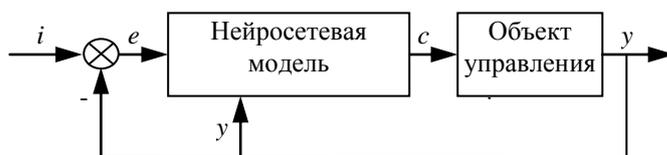


Рис.2. Замкнутая система «самолёт-лётчик» с нейросетевой моделью характеристик управляющих действий лётчика (обозначения те же самые, что и на рис. 1)

Одна из важнейших задач, которую необходимо решить при построении требуемой

НС-модели, состоит в определении показателей соответствия между результатами НС-моделирования и полученными экспериментальными данными. В проведенных исследованиях были использованы следующие четыре показателя.

Первые два показателя вычисляются на основе значений управляющего сигнала $c_{экс}^*(t)$, формируемого летчиком в ходе эксперимента на рабочей станции в моменты времени $t_i, i=0, \dots, N$. Значение сигнала $c_{экс}^*(t)$ в момент времени t_i будем далее для сокращения записи обозначать как $c_{экс}^*(i)$. Все экспериментальные данные, доступные при решении задачи (т.е. множество полученных отсчетов $c_{экс}^*(i)$ сигнала $c_{экс}^*(t)$), разбиваются на два непересекающихся подмножества. Одно из них будем далее именовать обучающим набором, второе – тестовым набором, а элементы этих наборов – примерами (обучающими и тестовыми, соответственно).

Первый из показателей, который будем обозначать как $mse(c)_{training}$, требуется для обеспечения процесса обучения НС-модели (см. рис.3). Он определяется на обучающем наборе и оценивает среднеквадратическую ошибку воспроизведения НС-моделью данного набора, состоящего из N примеров $c_{экс}^*(i)$.

Данный показатель оценивает степень различия между сигналом $c_{экс}^*(t)$ и сигналом $c_{мод}^*(t)$, выдаваемым НС-моделью. Он формулируется следующим образом:

$$mse(c)_{training} = \frac{1}{N} \sum_{i=1}^N [c_{экс}^*(i) - c_{мод}^*(i)]^2 \quad (1)$$

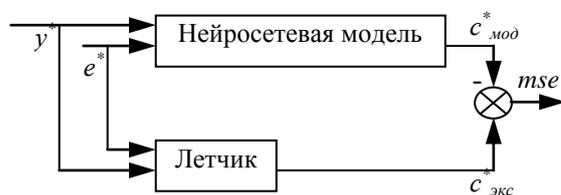


Рис.3. Сопоставление управляющих действий летчика и НС-модели при формировании показателя (1)

Если значение показателя (1) стремится к нулю в ходе процесса обучения НС-модели, то это означает, что реакция $c_{мод}^*(t)$ полученной модели близка к реакции $c_{экс}^*(t)$ летчика и,

следовательно, НС-модель хорошо аппроксимирует характеристики управляющих действий летчика.

Второй показатель, который будем обозначать как $mse(c)_{checking}$, имеет точно такой же вид, что и показатель (1):

$$mse(c)_{checking} = \frac{1}{N} \sum_{i=1}^N [c_{экс}(i) - c_{мод}(i)]^2 \quad (2)$$

Разница между показателями (1) и (2) состоит в характере данных, используемых для их подсчета. Если показатель (1) рассчитывался на обучающем наборе, то показатель (2) вычисляется на тестовом наборе. Смысл вычисления показателя (2) состоит в следующем. Сформированная НС-модель может хорошо (в указанном выше смысле) воспроизводить значения обучающего набора, однако при значениях сигнала $c_{экс}^*(t)$, отличающихся от обучающих примеров, погрешность модели может недопустимо возрасти. Чтобы удостовериться в отсутствии такого рода эффекта и производится вычисление показателя (2), которое осуществляется на данных, заведомо не использовавшихся при обучении НС-модели. Если на этих данных значение показателя $mse(c)_{checking}$ стремится к нулю в ходе процесса обучения, это свидетельствует о том, что жесткой привязки НС-модели к обучающим данным нет, то есть эта модель в состоянии воспроизводить с требуемой точностью и примеры, не входившие в состав обучающего набора.

Третий и четвертый показатели используются для оценки работы обученной НС-модели в замкнутой системе. Третий показатель позволяет найти среднеквадратическую ошибку воспроизведения НС-моделью тестового набора $y_{экс}(t)$ во временной области:

$$mse(y)_{closeloop} = \frac{1}{N} \sum_{i=1}^N [y_{экс}(i) - y_{мод}(i)]^2 \quad (3)$$

Четвертый показатель определен в частотной области и вычисляется следующим образом:

$$J = \sum_{k=1}^{15} [A_{экс}(\omega_k) - A_{мод}(\omega_k)]^2 + \frac{\pi}{180} [\varphi_{экс}(\omega_k) - \varphi_{мод}(\omega_k)]^2 \quad (4)$$

Здесь $A(\omega)$ и $\varphi(\omega)$ – амплитудные и фазовые частотные характеристики управляющих действий летчика, ω – частота. Этот показатель оценивает адекватность результатов моделирования в

частотной области и часто используется на практике.

Показатели, сформулированные выше, определяют требования к поведению НС-модели, другими словами, они фактически определяют характер ее работы. Для непосредственного формирования такого рода моделей, а также для проведения с ними соответствующих вычислительных экспериментов использовалась система Matlab и входящий в ее состав пакет Neural Networks. Обучение НС-модели осуществлялось с помощью обычного метода обратного распространения ошибки [5,6], основанного на алгоритме Левенберга-Марквардта, который показал себя наиболее эффективным и работоспособным среди набора алгоритмов обучения из состава средств упомянутого пакета.

В следующих разделах рассматриваются основные вопросы, связанные с подготовкой данных, требуемых для обучения НС-модели, с выбором структуры нейросети, а также с ее обучением. Был сформирован ряд НС-моделей, особенности которых анализируются ниже.

2. Формирование и обучение нейросетевой модели управляющих действий летчика

2.1. Формирование обучающего набора

Формирование обучающего набора имеет очень большое значение, поскольку он призван отражать основные закономерности управляющих действий летчика, которые и должна воспроизводить НС-модель. В задачах высокоточного пилотирования каждый из измеряемых сигналов в системе «самолет-летчик» (рис.4) можно представить в виде суммы коррелированной и некоррелированной с входным сигналом составляющих:

$$e = e_i + e_n, \quad c = c_i + c_n, \quad y = y_i + y_n$$

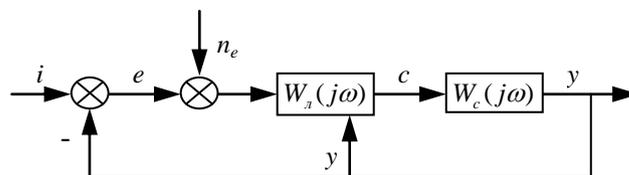


Рис. 4. Сигналы в системе «самолёт-лётчик» в задачах экстремального высокоточного пилотирования

Разработанные в [1] методы измерений позволяют как определить полные сигналы e , c , y так и

выделить из них коррелированные составляющие e_i , c_i , y_i . Сигналы e_n , c_n , y_n определяют действие на систему шума, вводимого летчиком.

В качестве обучающего набора для построения НС-модели принимается набор значений коррелированной с входным воздействием составляющей реакции летчика, которая определяет соответствующую составляющую ошибки слежения. Она вычисляется с помощью следующих соотношений:

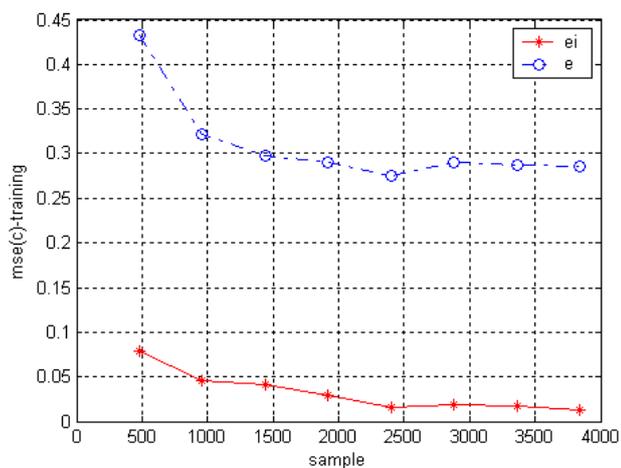
$$c_i = \sum_{k=1}^{15} \left| \frac{\phi}{W_c} \right| A_k \sin(\omega_k t + \varphi_{\frac{\phi}{W_c}})$$

$$y_i = \sum_{k=1}^{15} |\phi| A_k \sin(\omega_k t + \varphi_{\phi})$$

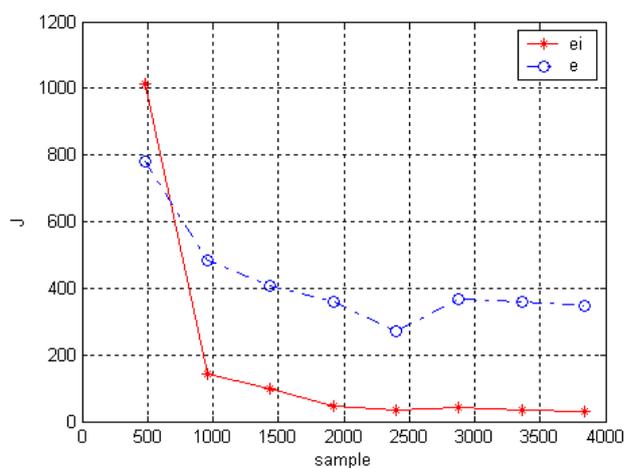
$$e_i = i - y_i$$

где ϕ – передаточная функция замкнутой системы, $i = \sum_{k=1}^{15} A_k \sin(\omega_k t)$.

Объем обучающего набора существенно влияет на результаты и трудоемкость обучения НС-модели. Небольшой объем обучающего набора может оказаться недостаточным для обучения с заданным уровнем точности. Напротив, слишком большой объем обучающего набора не улучшает существенно результаты обучения, но значительно увеличивает его трудоемкость. По этой причине в работе было проведено исследование рационального объема используемых экспериментальных данных.



(a)



(б)

Рис.5. Значения показателей $mse(c)_{training}$ и J для различных объемов обучающего набора

На рис.5 приведены данные, позволяющие судить о точности получаемых НС-моделей по показателям $mse(c)_{training}$ и J для различных объемов обучающего набора. Полученные результаты показывают, что в случае НС-модели с линейной активационной функцией результаты обучения для обучающего набора без шума (случай e_i) будут лучше, чем для обучающего набора с шумом (случай e). С ростом объема обучающего набора вначале оба показателя быстро уменьшаются, затем это уменьшение резко замедляется. Когда объем обучающего набора составляет более 2400 точек (обучающих примеров), точность НС-модели изменяется незначительно. В связи с этим, дальнейшим был выбран сигнал e_i , некоррелированный с шумом, с объемом обучающего набора, равным 2400 точек.

2.2. Определение структуры нейросетевой модели

Определение структуры нейросетевой модели включает формирование набора входов и выходов, определение числа скрытых слоев и числа нейронов в этих слоях, выбор активационной функции нейронов.

Формирование набора входов и выходов. На рис.6 показана НС-модель самой простой структуры, в которой всего один нейрон, имеющий линейную активационную функцию. На рис.7 представлен вид такой активационной функции.

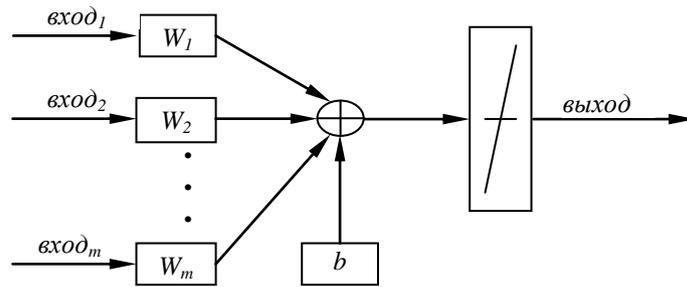
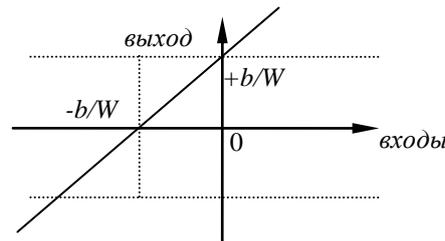


Рис.6. Структура нейросетевой модели с одним нейроном с линейной активационной функцией



$$\text{выход} = \text{purelin}(W \times \text{входы} + b) = W \times \text{входы} + b$$

Рис.7. Линейная активационная функция

Как уже отмечалось выше, выход НС-модели – сигнал c , а входы НС-модели состоят из сигнала e , производных сигнала e , сигнала y и производных сигнала y . Данные о производных сигналах e и y непосредственно не используются, вместо этого учитывается предыстория соответствующих сигналов глубиной в несколько временных шагов.

Для построения соответствующей НС-модели использовалась сеть с архитектурой типа TDNN (Time Delay Neural Network). Построенная на основе персептрона и линий задержки для входных сигналов, она позволяет учитывать не только значение входного вектора для текущего момента времени, но и предысторию изменения этого значения. Для данной архитектуры в компьютерном эксперименте варьировались количество входов сети (на этот параметр влияет, прежде всего, используемая глубина предыстории для входных сигналов), величина задержек на входе модели и шаг, который используется при вычислении производной.

Чем больше входов поступает в сеть (за счет увеличения глубины предыстории), тем больше варьируемых параметров в НС-модели, что дает возможность повысить точность данной модели, оцениваемую показателем $mse(c)_{training}$. Однако слишком большая точность по этому

показателю, достигнутая в процессе обучения нейросети, ведет к тому, что модель начинает запоминать и воспроизводить такие подробности обучающего набора, которые не должны присутствовать в получаемой модели. В итоге полученная модель оказывается неудовлетворительной по своим характеристикам в составе замкнутой системы, для которой она и создается. В этом случае можно говорить о так называемой переразмеренности нейросетевой модели (в ней слишком много параметров, число «степеней свободы» в ней слишком велико, она получается излишне чувствительной по отношению к особенностям обучающего набора). Поэтому число входов НС-модели следует ограничить. При этом следует стремиться к тому, чтобы с достаточной точностью описать характеристики управляющих действий летчика с помощью модели, имеющей наиболее простую структуру.

На рис.8 показаны результаты обучения НС-модели по показателям $mse(c)_{training}$ и J для различных значений числа входов e и Δe , а также при различных значениях задержки ($\tau_e = \tau_{\Delta e}$) при формировании предыстории.

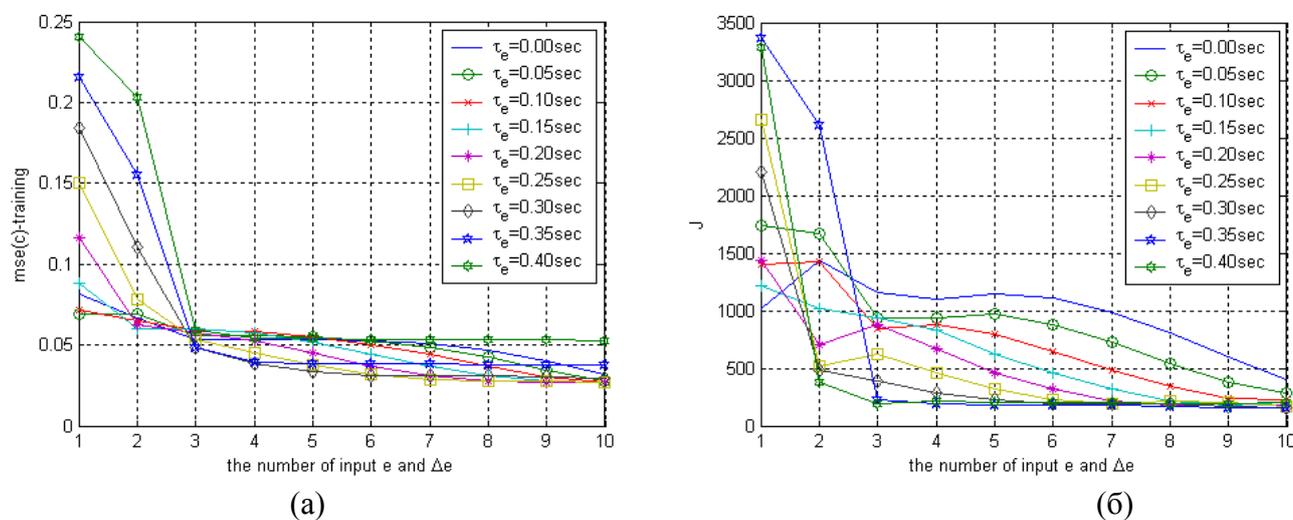


Рис.8. Результаты обучения НС-модели для различных значений числа входов e и Δe , а также при различных значениях задержки ($\tau_e = \tau_{\Delta e}$) при формировании предыстории: **(а)** по показателю $mse(c)_{training}$; **(б)** по показателю J

Из этих результатов видно, что увеличение числа входов до трех ведет к существенному уменьшению значений показателей $mse(c)_{training}$ и J . Это значит, что число входов не следует

принимать меньшим 3. Увеличение числа входов сверх 3 слабо сказывается на значениях показателей $mse(c)_{training}$ и J . По этой причине величина числа входов для сигнала e и Δe была выбрана равной 3. Чем больше задержки сигналов e и Δe , тем меньше значения показателя J . Наилучшие значения по этим показателям получаются при числе входов, равном 3 и величине задержек $\tau_e = \tau_{\Delta e} = 0.35$ или 0.40.

Аналогичным образом выбиралось число входов для сигнала Δy (см. рис.9).

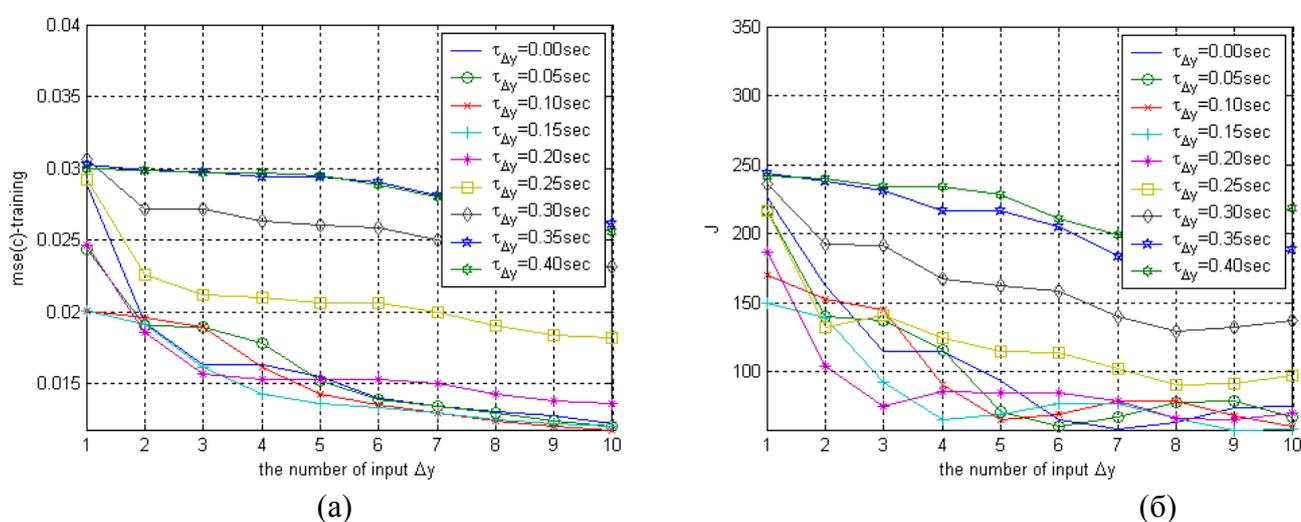
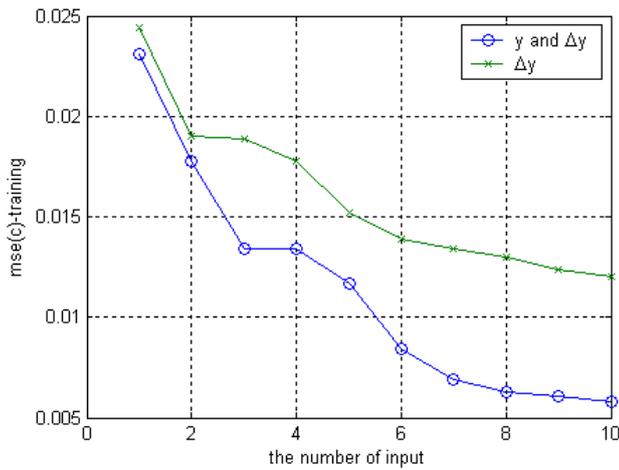


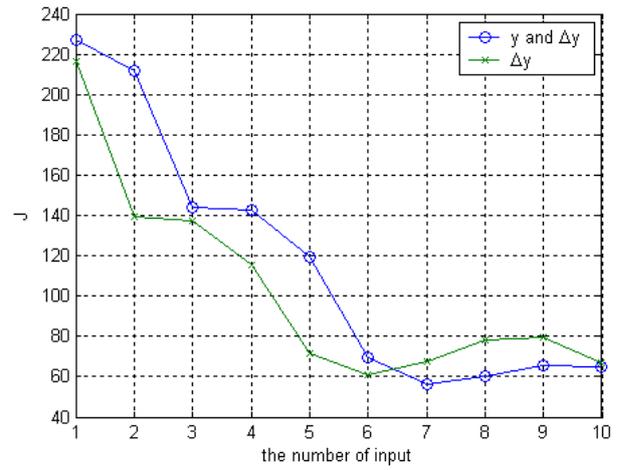
Рис.9. Результаты обучения НС-модели для различных значений числа входов Δy , а также при различных значениях задержки ($\tau_e = \tau_{\Delta e}$) при формировании предыстории: (а) по показателю $mse(c)_{training}$; (б) по показателю J

При больших значениях задержек сигнала Δy показатели $mse(c)_{training}$ и J также принимают большие значения. Задержки сигнала Δy нельзя принимать большими, чем 0.25. Как и в случае сигналов e и Δe , число входов для сигнала Δy было выбрано равным 3.

Поскольку скорость изменения сигнала y влияет на управление, реализуемое летчиком, следует проверить, насколько необходим входной сигнал y . Было проведено сравнение вариантов входного набора с сигналом y и без него. На рис.10 показаны результаты обучения НС-модели по показателям $mse(c)_{training}$ и J для этих двух случаев.



(a)



(б)

Рис.10. Результаты обучения НС-модели с сигналом y и без него: **(а)** по показателю $mse(c)_{training}$; **(б)** по показателю J

Полученные результаты показывают, что по показателю $mse(c)_{training}$ использование сигнала y в качестве входного дает существенное уменьшение получаемой ошибки, однако по показателю J , то есть для случая замкнутой системы, который является основным, картина обратная – введение сигнала y в качестве входного дает увеличение ошибки, когда количество входов меньше, чем 6. В связи с этим сигнал y из числа входных в данном случае следует исключить.

Входы сигналов e , Δe и Δy имеют задержки, равные $\tau_e, \tau_{\Delta e}$ и $\tau_{\Delta y}$, соответственно. При этом сигналы Δe и Δy изменяются с шагом, равным $\delta_{\Delta e}$ и $\delta_{\Delta y}$, соответственно. Ранее был получен результат, согласно которому $\tau_e = \tau_{\Delta e} = 0.35$ или 0.40 с. Однако значения задержек τ_e и $\tau_{\Delta e}$ не могут быть одинаковыми. Согласно смыслу решаемой задачи, параметр τ_e означает задержку реакции летчика на сигнал e . Ее величина составляет примерно $\tau_e = 0.20$ с. А значение параметра $\tau_{\Delta e}$ состоит из той же самой задержки реакции летчика на сигнал e плюс время вычисления скорости изменения сигнала e , т.е. оно должно быть больше, чем значение параметра τ_e . На рис.11 показаны результаты обучения НС-модели по показателю J для разных значений задержек для e и Δe . Результаты показывают, что по показателю J малые значения его были получены в трех случаях: при $\tau_e = 0.10$ с и $\tau_{\Delta e} = 0.20$ с; при $\tau_e = 0.25$ с и $\tau_{\Delta e} = 0.40$ с; при $\tau_e = 0.35$ с и $\tau_{\Delta e} = 0.35$ с. При τ_e

$\tau_e=0.10$ с значение показателя получается слишком маленьким, а при $\tau_e=0.35$ с – слишком большим.

Для формирования НС-модели был выбран второй из перечисленных выше случаев.

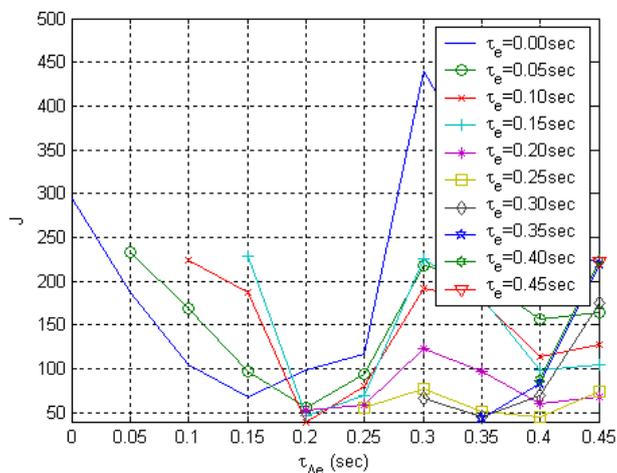


Рис.11. Результаты обучения НС-модели по показателю J для разных значений задержек e и Δe

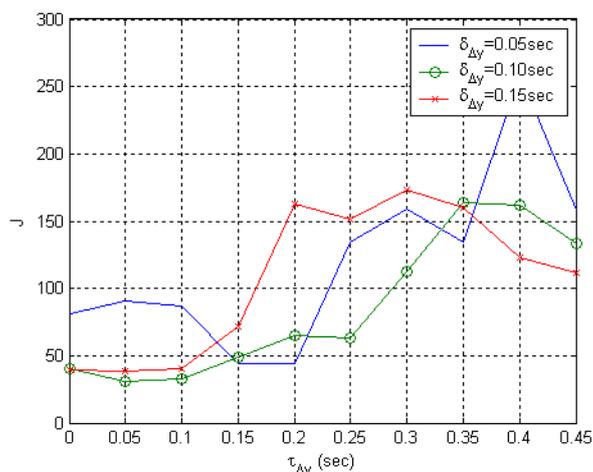


Рис.12. Результаты обучения НС-модели по показателю J для разных значений задержек Δy и шага Δy

На рис.12 показаны результаты обучения НС-модели по показателю J для разных значений задержек Δy и шага Δy . Здесь $\tau_{\Delta y}=0.05$ с, а $\delta_{\Delta y}=0.10$ с – самое маленькое значение. Значение $\delta_{\Delta e}$ выбрано равным 0.05 с, поскольку сигнал e изменяется медленнее сигнала y .

На рис.13 показаны входы и выходы НС-модели с учетом полученных результатов.

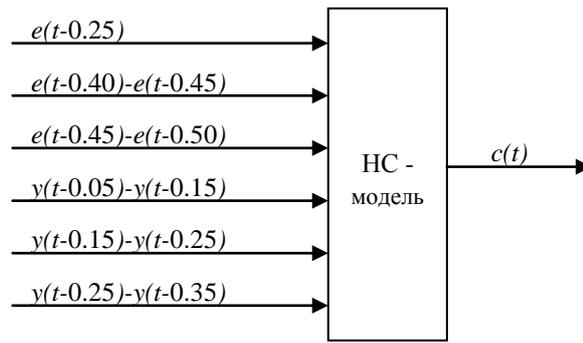


Рис.13. Входы и выходы полученной НС-модели

На рис.14 показаны результаты моделирования и частотные характеристики летчика. Видно, что НС-модель достаточно хорошо отражает характеристики действий летчика в задачах компенсаторного слежения.

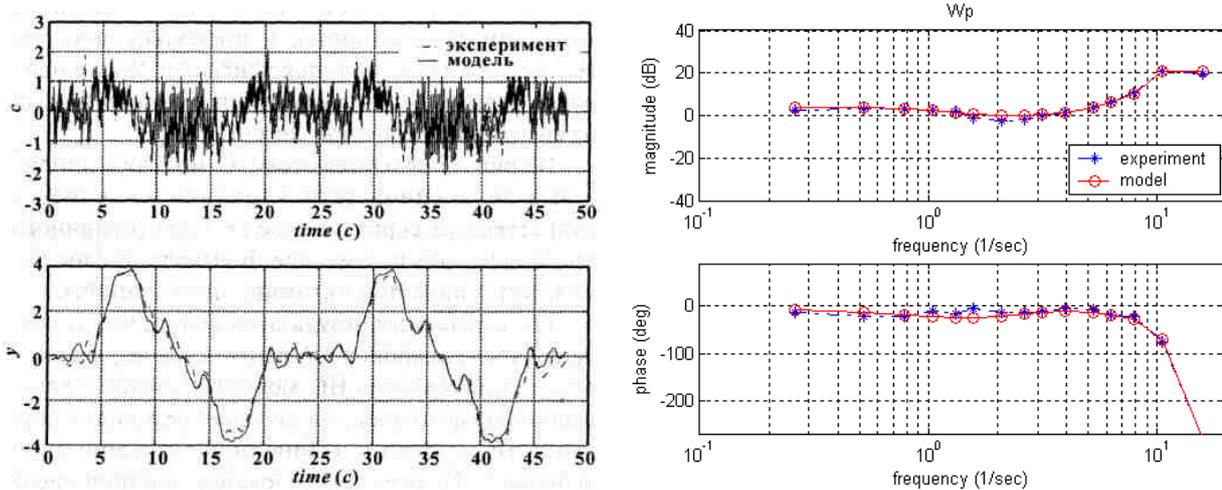


Рис. 14. Результаты моделирования и частотные характеристики летчика

Определение числа скрытых слоев и числа нейронов в них, выбор активационной функции нейрона. В исследованиях, описанных выше, была использована линейная НС-модель с единственным нейроном, который имеет линейную активационную функцию. В данном разделе для исследования характеристик действий летчика НС-модель строится на основе мультиперсептрона, в котором может варьироваться число скрытых слоев, число нейронов в каждом слое и вид активационной функции в нейронах скрытого слоя. Для формируемой НС-модели целесообразно выбрать структуру с одним скрытым слоем [5, 6] (рис.15).

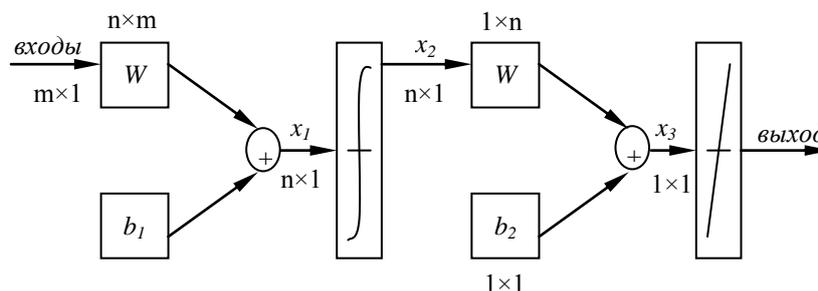


Рис.15. Структура нейросетевой модели с одним скрытым слоем

В нейронах скрытого слоя используются нелинейные активационные функции-сигмоиды (логистическая функция `logsig` пакета `Neural Networks`), показанные на рис.16.

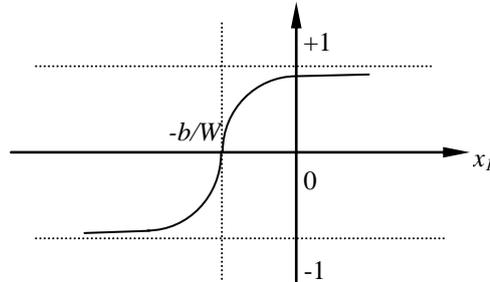
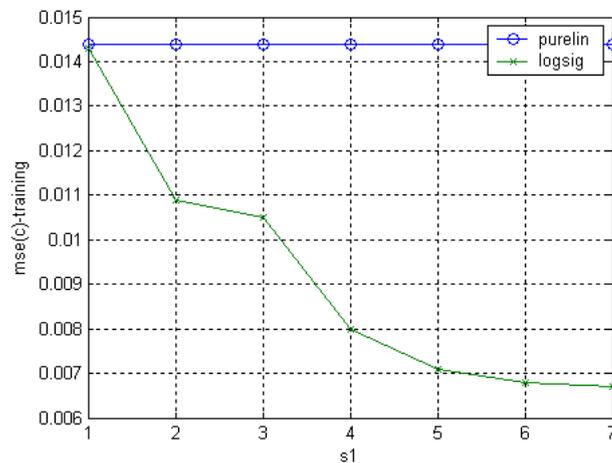


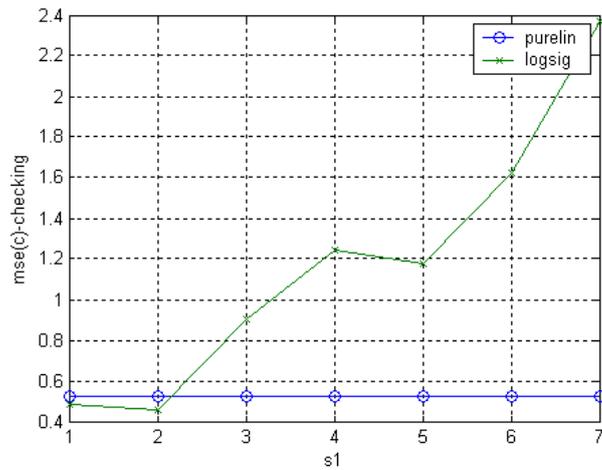
Рис.16. Логистическая активационная функция нейрона (сигмоида)

В выходном слое сохранена линейная активационная функция (`purelin`). Число нейронов в выходном слое равняется 1, поскольку выходом НС-модели является только сигнал c . Число нейронов скрытого слоя требуется определить, исходя из специфики решаемой задачи.

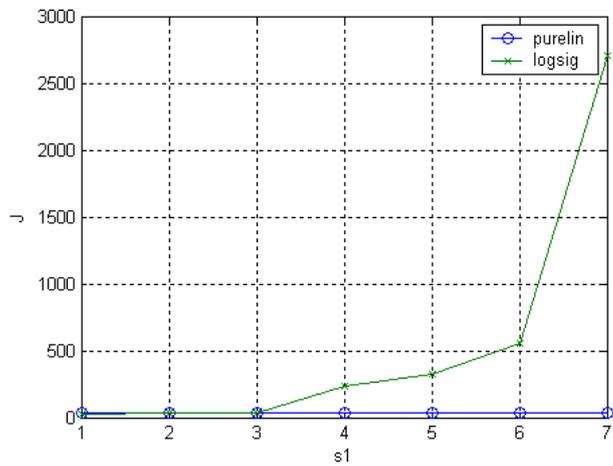
На рис.17 сравниваются НС-модели с линейной и нелинейной активационной функцией в единственном скрытом слое ($s1$) для различного числа нейронов в этом слое. В выходном слое используется линейная активационная функция.



(a)



(б)



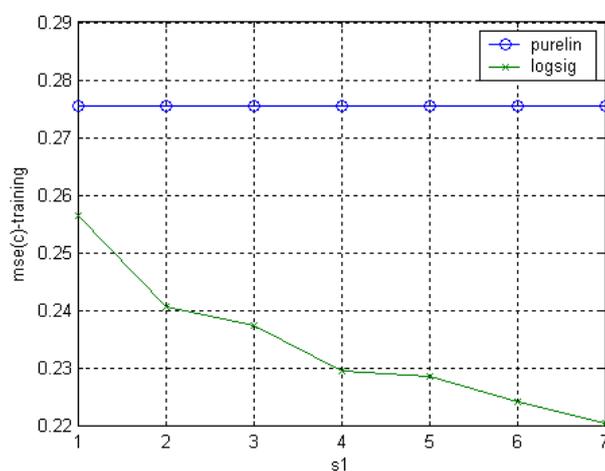
(в)

Рис.17. Сравнение НС-моделей с линейной и нелинейной активационной функцией в единственном скрытом слое для различного числа нейронов (s1) в этом слое: **(а)** результаты по показателю $mse(c)_{training}$; **(б)** – по показателю $mse(c)_{checking}$; **(в)** – по показателю J .

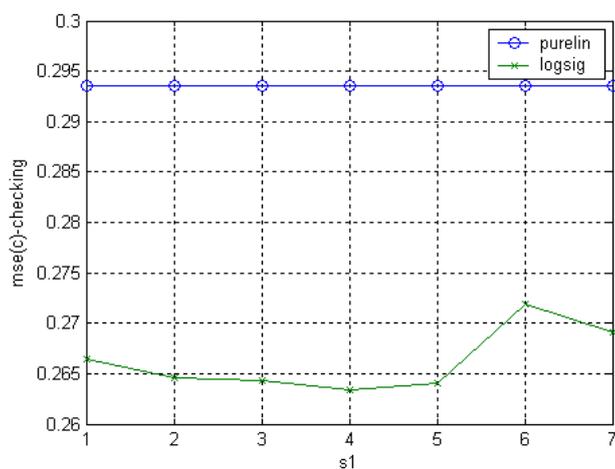
Из полученных результатов видно, что увеличение числа нейронов в скрытом слое не улучшает результатов обучения НС-модели с линейной активационной функцией, но улучшает результаты обучения НС-модели с нелинейной активационной функцией. Поэтому исследование влияния числа нейронов в скрытом слое проводится в дальнейшем только для НС-модели с нелинейной активационной функцией. НС-модель с нелинейной активационной функцией позволяет получить меньшие значения показателя $mse(c)_{training}$, чем НС-модели с линейной активационной функцией. Это значит, что НС-модель с нелинейной активационной функцией обладает обучаемостью. Однако для показателя $mse(c)_{checking}$ был получен результат

противоположного характера. Отсюда следует, что НС-модель с нелинейной активационной функцией менее работоспособна при решении анализируемой задачи. Поскольку из обучающего набора без шума убрана нелинейная составная часть входной информации, нелинейная активационная функция уже не отвечает рассматриваемому случаю. Результаты моделирования показывают, что при малых значениях показателя $mse(c)_{training}$, показатель J может достигать значительной величины. Это связано с проблемой переобучения нейросети, которая будет рассмотрена ниже.

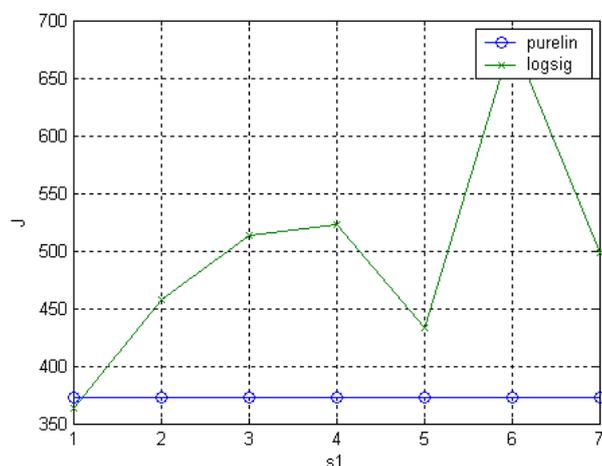
На рис.18 показаны результаты сравнения НС-моделей с линейной и нелинейной активационной функцией для обучающего набора с шумом. По показателю $mse_{checking}$ НС-модель с нелинейной активационной функцией хуже, чем НС-модель с линейной активационной функцией. По показателю J НС-модель с нелинейной активационной функцией хуже, чем НС-модель с линейной активационной функцией. По-видимому, показатель J неадекватен задаче описания нелинейных характеристик. В связи с этим при оценке НС-модели с нелинейной активационной функцией следует использовать показатель $mse_{checking}$. Для случая замкнутой системы требуется формирование нового критерия.



(a)



(б)



(в)

Рис.18. Сравнение НС-моделей с линейной и нелинейной активационной функцией для обучающего набора с шумом: (а) результаты по показателю $mse(c)_{training}$; (б) – по показателю $mse(c)_{checking}$; (в) – по показателю J .

Согласно полученным результатам, число нейронов в скрытом слое НС-модели требовалось выбрать меньшим трех, например, его можно принять равным двум.

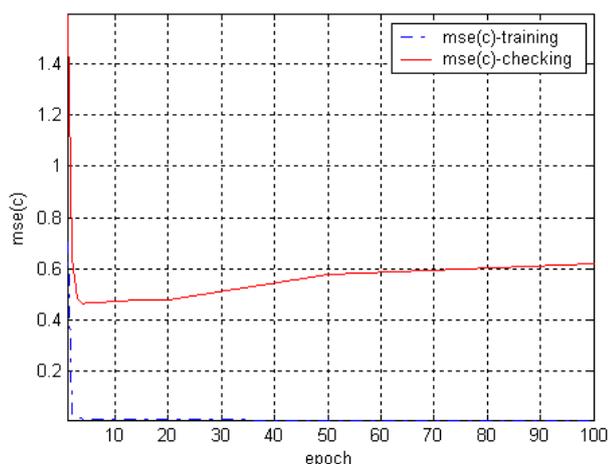
2.3. Проблема переобученности нейросетевой модели

Проведенные исследования выявили существенное влияние избыточного обучения (переобучения) НС-модели на ее адекватность решаемой задаче управления и соответствие поведения модели экспериментальным данным. Решение проблемы переобучения представляет собой процесс, состоящий из последовательности этапов, среди которых: выбор структуры НС-модели; определение методики обучения НС-модели, включая выявление рационального

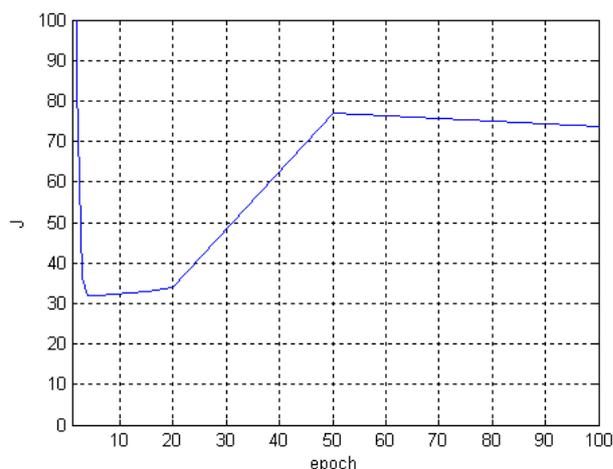
объема используемых экспериментальных данных, а также величины шага их представления; оценка влияния указанных выше факторов на процесс обучения НС-модели и на ее работоспособность при расширении объема используемых обучающих данных.

Выше была выявлена рациональная структура НС-модели, а также требуемый состав обучающего набора, основанного на экспериментальных данных. В качестве метода обучения во всех случаях, рассматриваемых в данной работе, используется метод обратного распространения ошибки. Теперь необходимо обсудить влияние на характеристики получаемой НС-модели продолжительности процесса обучения, измеряемой в числе выполненных итераций («эпох»), каждая из которых состоит в предъявлении всех примеров из обучающего набора и вычислении суммарной среднеквадратичной ошибки.

После того, как НС-модель с нелинейной активационной функцией была обучена на наборе без шума для 1000 итераций в процессе обучения НС-модели, эта модель была использована в замкнутой системе с вычислением соответствующих частотных характеристик. Результаты показали неудовлетворительную точность, получаемую в данном случае. На рис.19 приведены значения показателей $mse(c)_{training}$, $mse(c)_{checking}$ и J для различных значений числа итераций в процессе обучения НС-модели. Когда число итераций равняется 4, значения $mse(c)_{checking}$ и J минимальны. Если число итераций превышает 4, значения показателей $mse(c)_{checking}$ увеличиваются. Это значит, что работоспособность НС-модель в этом случае ухудшится. Из полученных результатов следует существенная зависимость характеристик получаемой НС-модели от числа итераций процесса обучения. Это число должно быть строго ограниченным, в противном случае НС-модель теряет свою адекватность в замкнутой системе.



(a)



(б)

Рис.19. Значения показателей $mse(c)_{training}$, $mse(c)_{checking}$ и J для различных значений числа итераций в процессе обучения НС-модели

Сравнение показателей для НС-моделей с линейными и нелинейными активационными функциями, а также для обучающих наборов с шумом и без шума приведено в табл. 1.

Таблица 1

Структура	Обучающий набор	$mse(c)_{training}$	$mse(c)_{checking}$	J
S1=1,func1=purelin	без шума	0.0144	0.5225	30.459
S1=2,func1=logsig	без шума	0.0146	0.4627	32.070
S2=1,func2=purelin	с шумом	0.2573	0.2315	462.55

Сравнение случаев для обучающего набора без шума показали, что значения показателей $mse(c)_{training}$ для них близки. По показателю $mse_{checking}$ НС-модели с нелинейной активационной функцией точнее. Однако для показателя J точнее НС-модели с линейной активационной функцией. Сравнение случаев для нелинейных активационных функций показывает, что значения показателей $mse(c)_{training}$ и J для обучающего набора с шумом больше, чем для обучающего набора без шума. Но для показателя $mse_{checking}$ значение для обучающего набора с шумом меньше, чем для обучающего набора без шума.

Итак, для линейного объекта НС-модель с линейной активационной функцией достаточно хорошо описывает характеристики действий летчика. НС-модели с нелинейной активационной функцией позволяют получить несколько лучшие результаты. Однако структура такой модели

сложнее и она подвержена переобучению. НС-модель с нелинейной активационной функцией следует применять лишь там, где требуется описывать нелинейные свойства характеристик управляющих действий летчика. Этот случай в дальнейшем не рассматривается, поскольку в данной работе такого рода нелинейность не учитывается.

2.4. Сравнение нейросетевой модели и традиционных моделей

Работоспособность сформированной НС-модели проверялась в замкнутом контуре системы «самолет-летчик». Полученные результаты показали более высокую степень согласованности с экспериментальными данными, чем это имеет место для широко распространенных квазилинейных моделей характеристик управляющих действий летчика (рис.20).

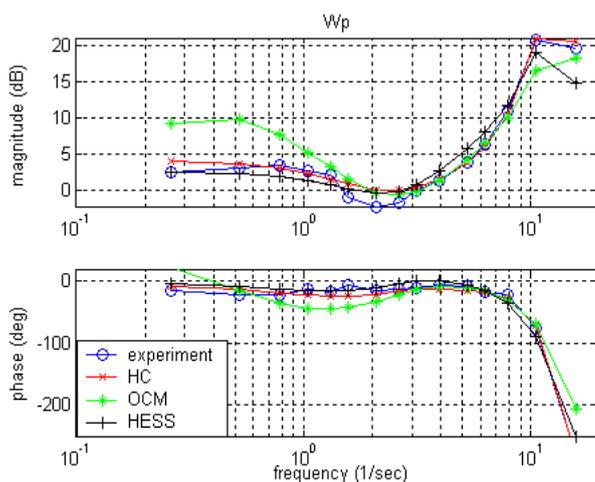


Рис.20. Степень согласованности с экспериментальными данными для нейросетевой модели и традиционных моделей характеристик управляющих действий летчика

3. Адаптируемость НС-модели

Представленная выше серия экспериментов была связана с получением НС-модели управляющих действий летчика для объекта управления следующего вида:

$$W_c = \frac{1 + T_\theta s}{s(1 + 2\xi Ts + T^2 s^2)}$$

Здесь было принято, что $T_\theta = 1.4$, $\xi = 0.64$, $1/T = 2.4$ и $K_c = 1.5$. Однако хорошая НС-модель должна не только описывать характеристики управляющих действий летчика применительно к одному объекту управления, но и сохранять адекватность в случае перехода к другому объекту. В данном

разделе приводятся результаты экспериментов, позволяющие оценить степень адаптивности НС-модели, упомянутой выше, при переходе к другим объектам управления.

Эксперименты проводились для 15 разных объектов, у которых $1/T=1.6, 2.0$ или 2.4 , а $K_c = 0.5, 1.5, 2.5, 3.5$ или 4.5 . Эти объекты обозначаются как 1605, 1615,2445, где первые две цифры означают величину, обратную постоянной времени, т.е. $1/T$, а последние две – значение коэффициента усиления K_c . Для каждого из объектов летчиком выполнялся эксперимент. На рис.21 показаны дисперсии сигналов e и e_i (σ_e^2 и $\sigma_{e_i}^2$) для каждого эксперимента. Черным здесь показаны величины $\sigma_{e_i}^2$, а белым – величины $\sigma_e^2 - \sigma_{e_i}^2$ (шум).

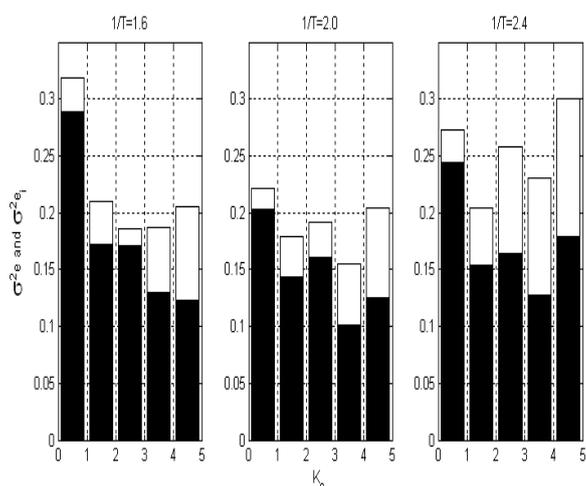


Рис.21. Дисперсии сигналов e и e_i (σ_e^2 и $\sigma_{e_i}^2$) для экспериментов с различными моделями объектов управления: черным показана величина $\sigma_{e_i}^2$, белым – величина $\sigma_e^2 - \sigma_{e_i}^2$ (шум).

Полученные результаты показывают, что при больших значениях $1/T$ ухудшаются характеристики управляющих действий летчика (значение σ_e^2 для случая $1/T=2.4$ больше, чем для случая $1/T=1.6$ и 2.0). Маленькие или большие значения K_c также ухудшают характеристики управляющих действий летчика (значение σ_e^2 для случая $K_c=0.5$ больше, чем в других случаях, значение шума для случая $K_c=3.5$ или 4.5 также больше, чем в других случаях).

Согласно полученной выше структуре было осуществлено обучение НС-модели. На рис.22 показаны результаты, показывающие, что чем больше значение коэффициента K_c , тем больше

получаемое значение показателя $mse(y)_{closeloop}$. По показателю J все результаты удовлетворительны ($J < 200$), кроме случая при $K_c=4.5$.

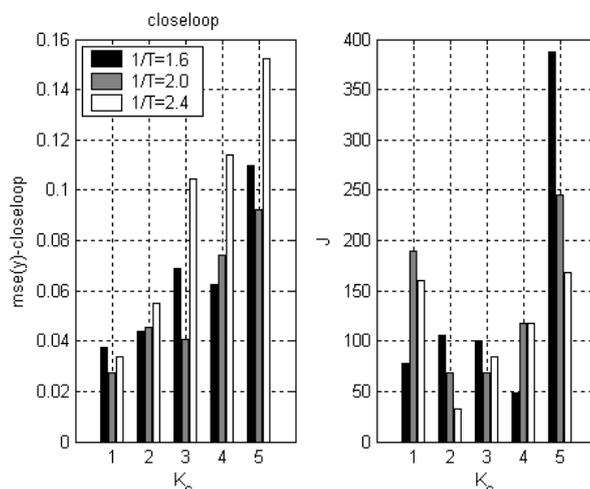


Рис.22. Значения показателей $mse(y)_{closeloop}$ и J в зависимости от величины коэффициента усиления K_c при различных значениях параметра $1/T$

4. Прогнозирующие возможности НС-модели

Математическая модель характеристик управляющих действий летчика должна сохранять адекватность в случае, когда изменяется объект управления или входной сигнал. Такого рода ее свойство будем называть прогнозирующей способностью модели. Очевидно, что это требование распространяется и на формируемые НС-модели. Результаты, приводимые ниже, дают возможность сделать определенные выводы относительно прогнозирующих свойств НС-моделей.

4.1. Влияние коэффициента усиления

Применительно к объектам 2405, 2415, 2425, 2435 и 2445 была проведена серия компьютерных экспериментов, в которой НС-модель, построенная для одного объекта управления, использовалась для прогнозирования характеристик управляющих действий летчика для другого объекта. Полученные результаты показывают, что если разница в значениях коэффициентов усиления для пары рассматриваемых объектов достаточно велика, то в замкнутом контуре, включающем построенную модель, наблюдается расходимость процессов. Поскольку произведение коэффициентов усиления летчика и объекта управления есть постоянная величина, можно изменять коэффициенты НС-модели так, чтобы обеспечить адекватное прогнозирование

характеристик управляющих действий летчика для объекта, отличающегося от того, для которого формировалась НС-модель:

$$W_{\text{проверка}} = W_{\text{обучение}} \times K_{\text{с-обкчение}} \cdot K_{\text{с-проверка}}$$

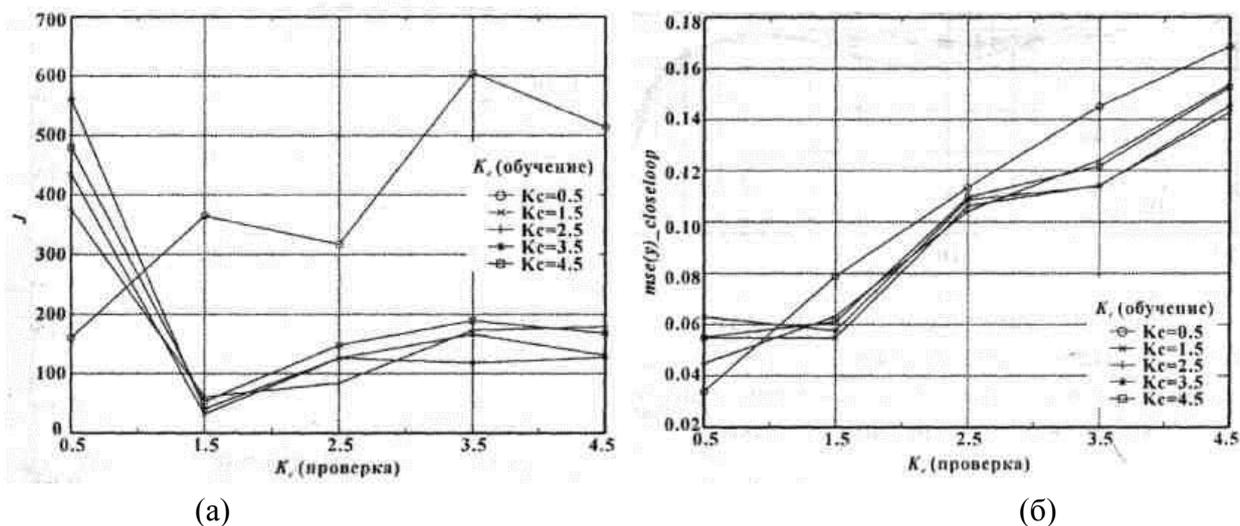


Рис.23. Результаты проверки адекватности НС-модели управляющих действий летчика, построенной для одного объекта управления, на других объектах: **(а)** по показателю $mse(y)_{\text{checking}}$; **(б)** по показателю J

Полученные результаты показаны на рис.23. Из них видно, что кроме случая $K_c = 0.5$, получаются хорошие результаты прогноза ($J < 200$). Оказалось, что построение НС-модели, описывающей характеристики управляющих действий летчика для объекта, имеющего маленький коэффициент усиления, сталкивается с определенными трудностями.

Можно также для прогнозирования характеристик управляющих действий летчика воспользоваться интерполяцией коэффициентов НС-модели (ее синаптических весов). Сравнение результатов применительно к разным вариантам прогноза для разных значений коэффициента усиления показаны на рис.24. Видно, что результаты прогноза для варианта с корректировкой коэффициентов НС-модели и интерполяционного варианта близки. Значения показателя J для варианта без корректировок коэффициента усиления слишком велики в области малых значений K_c . Интерполяционный вариант, а также вариант с прямым обучением требуют для своей реализации достаточного количества моделей для разных объектов. В связи с этим наиболее целесообразно использовать вариант с корректировкой коэффициента усиления.

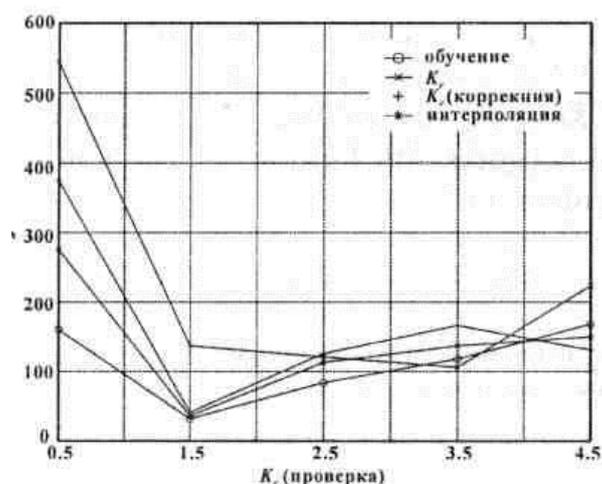


Рис.24. Результаты для разных вариантов прогноза в зависимости от величины коэффициента усиления

4.2. Влияние временных задержек

Если объект управления имеет большую по величине постоянную времени, характеристики управляющих действий летчика должны иметь малую временную задержку. Величину задержки для НС-модели летчика следует варьировать с помощью фильтра сигнала $\Delta y - 1/(T_y s + 1)$. Для больших значений постоянной времени объекта надо выбирать малые значения параметра T_y . Применительно к объектам 1615, 2015 и 2415 ($K_c=1,5$ и $1/T=1,6; 2,0; 2,4$) была проведена серия компьютерных экспериментов, в которой НС-модель, построенная для одного объекта управления, использовалась для прогнозирования характеристик управляющих действий летчика для другого объекта. Для разных значений частоты колебаний объекта ($\omega=1/T$) были выбраны разные значения T_y (табл.2).

Таблица 2

$1/T$	1.6	2.0	2.4
T_y	0.2 с	0.1 с	0.0 с

На рис.25 показаны результаты прогноза НС-модели по показателям $mse(y)_{closeloop}$ и J .

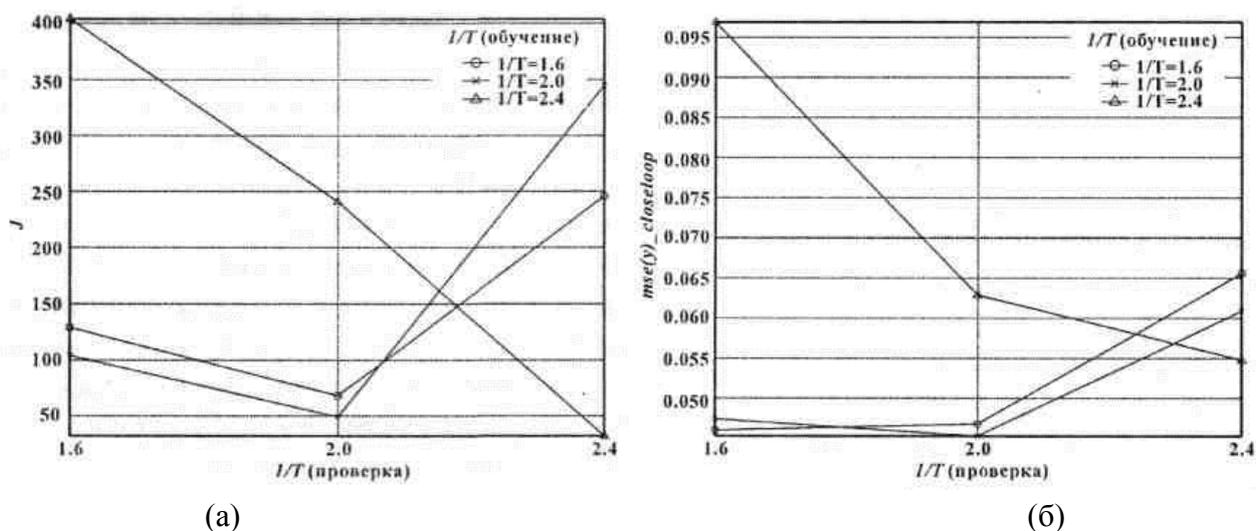


Рис.25. Результаты прогноза с помощью НС-модели по показателям $mse(y)_{close\ loop}$ и J для различных значений $1/T$ ($K_c=1.5$)

Выводы

Полученные в работе результаты позволяют сделать следующие выводы.

На характер построенной НС-модели влияет выбор состава ее входов. Определенное влияние оказывает также и вариативность обучающих наборов, полученных по результатам экспериментов с привлечением различных летчиков. Оценка этого влияния и возможности компенсации его за счет обобщающих свойств и адаптивности НС-модели требуют дополнительного исследования.

Для линейного объекта НС-модель с линейной активационной функцией достаточно хорошо описывает характеристики действий летчика. В частности, такая модель адекватно отражает характеристики действий летчика в задачах компенсаторного слежения. НС-модели с нелинейной активационной функцией позволяют получить несколько лучшие результаты. Однако структура таких модели сложнее и они подвержены переобучению.

НС-модели показали более высокую степень согласованности с экспериментальными данными, чем это имеет место для широко распространенных квазилинейных моделей характеристик управляющих действий летчика. Для улучшения их прогнозирующих возможностей требуется вводить коррекцию для коэффициента усиления НС-модели и регулировать значение коэффициента фильтра для сигнала Δy .

Summary

Mathematical models of pilot control behavior characteristics are very important for manual control problems in regard closed aircraft-pilot systems. Conventional models of such nature suffer from some drawbacks which make difficult to solve application problems mentioned above. An artificial neural network based model of pilot control behavior is suggested to overcome these drawbacks. The model is derived from experimental data obtained during pilot-aided tests using a special workstation meant for manual control studies.

Работа выполнена при поддержке РФФИ (проект № 05-08-18118).

Библиографический список

1. *Ефремов А.В., Оглоблин А.В., Предтеченский А.Н., Родченко В.В.* Летчик как динамическая система. – М.: Машиностроение, 1992. – 343 с
2. *McRuer D.T., Krendel E.S.* Mathematical models of human pilot behavior. – North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development, AGARDograph No.188, 1974.
3. *Efremov A.V., et al.* Investigation of pilot-induced oscillation tendency and prediction criteria development. – WL-96-310999 Wright Lab. USA, May 1996. pp.1-138.
4. *Hess R.A.* Unified theory for aircraft handling qualities and adverse aircraft-pilot coupling // Journal of Guidance, Control and Dynamics. – 1997. – Vol. 20, No. 6. – pp. 1141-1148.
5. *Krose B., van der Smagt P.* An introduction to neural networks. 8th Edition. – The University of Amsterdam, 1996. – 135 pp.
6. *Hristev R.M.* The ANN book. – Electronic Edition, 1998. – xv+374 pp.