

ЗАСТОСУВАННЯ МАШИННОГО НАВЧАННЯ ДЛЯ УТОЧНЕННЯ ЧИСЕЛЬНИХ МЕТЕОРОЛОГІЧНИХ ПРОГНОЗІВ

А.Ю. Дорошенко^{a[0000-0002-8435-1451]}, В.М. Шниг^{b[0000-0003-1055-7120]},

Р.В. Кушніренко^{b[0000-0002-1990-8727]}

^aІнститут програмних систем НАН України, 03187, м. Київ-187, проспект Академіка Глушкова, 40.

^bУкраїнський гідрометеорологічний інститут ДСНС України та НАН України, 03028, м. Київ, проспект Науки, 37.

Зроблено короткий огляд тенденцій розвитку чисельного прогнозування погоди, труднощів та природи їх виникнення, існуючих та перспективних шляхів їх подолання. Як перспективний підхід для збільшення точності прогнозу приземної температури повітря чисельної регіональної моделі COSMO запропоновано архітектуру нейронної мережі, яка дозволяє прогнозувати помилки власне прогнозів атмосферної моделі із подальшим їх коригуванням. Проведено експерименти з різними передісторіями помилок регіональної моделі. Визначено кількість епох, після збільшення якої мало місце так зване перенавчання мережі. Показано, що запропонована архітектура дає можливість досягти покращення прогнозу приземної температури повітря приблизно у 50 % випадків.

Ключові слова: нейронна мережа, чисельна модель, COSMO, прогноз, температура повітря.

Сделан краткий обзор тенденций развития численного прогнозирования погоды, трудностей и природы их возникновения, существующих и перспективных путей их преодоления. Как перспективный подход для увеличения точности прогноза приземной температуры воздуха численной региональной модели COSMO предложена архитектура нейронной сети, которая позволяет прогнозировать ошибки собственно прогнозов атмосферной модели с последующей их коррекцией. Были проведены эксперименты с различными предисториями ошибок региональной модели. Определено количество эпох, после увеличения которого имело место так называемое переобучение сети. Показано, что предложенная архитектура позволяет достичь улучшения прогноза приземной температуры воздуха примерно в 50 % случаев.

Ключевые слова: нейронная сеть, численная модель, COSMO, прогноз, температура воздуха.

This paper presents a brief overview of trends in numerical weather prediction, difficulties, and the nature of their occurrence, the existing and promising ways to overcome them. The neural network architecture is proposed as a promising approach to increase the accuracy of the 2m temperature forecast given by the COSMO regional model (a non-hydrostatic model designed particularly for various scientific applications; based on the thermo-hydrodynamical equations describing compressible flow in a moist atmosphere), which is used by the Ukrainian Hydrometeorological Institute of the SES of Ukraine and the NAS of Ukraine for both operational and research applications since July 2011. The calculated region consists of 209 grid points in the east-west direction and 101 grid points in the north-south direction with a grid interval equal to approximately 14km; the number of vertical layers is 50. The proposed architecture allows predicting errors of the atmospheric model forecasts with their further corrections. The approach allows correcting inaccuracies caused by physical processes that the regional numerical model does not consider or consider partially. Different histories of regional model errors are taken into consideration (1, 2, 3, 5, 7-day ones). The number of epochs after which network overfitting happens is determined. All experiments are conducted using Keras (an open-source library that provides a Python interface for artificial neural networks) and Tensorflow (an open-source machine learning platform). The code was written in the Python programming language. It is shown that the proposed architecture makes it possible to achieve an improvement of a 2m temperature forecast in approximately 50% of cases.

Key words: neural network, numerical model, COSMO, forecast, air temperature

Вступ

За останні 30 років досягнуто безперечного прогресу в розробці і оперативному застосуванні сучасних чисельних гідродинамічних методів прогнозу погоди. Це стало можливим передусім завдяки появі потужних комп'ютерних систем, розробці більш довершених і реалістичних математичних моделей, а також істотному збільшенню обсягу початкових метеорологічних даних.

Сучасні уявлення про закономірності перебігу основних фізичних і термодинамічних процесів у тропосфері, стратосфері й мезосфері та існуюча залежність між цими атмосферними процесами дає можливість описати їх рівняннями гідродинаміки [1–12], що, в свою чергу, дозволяє моделювати різні ланки загальної циркуляції атмосфери й прогнозувати погодні явища. Існуючі у світовій практиці прогностичні моделі класифікуються в залежності від типу метеорологічних процесів, від специфіки їхньої математичної реалізації, від терміну прогнозів, а також від розміру території, для якої вони складаються.

Сьогодні удосконалення якості прогнозів погоди відбувається за двома напрямками: поліпшення справджуваності прогнозу і збільшення його завчасності. Короткостроковий прогноз визначається переважно початковим станом атмосфери. Для середньострокових прогнозів потрібне як детальне знання початкових полів метеорологічних величин, так і уміння описати впливи зовнішніх чинників, які призводять до нового стану рівноваги системи, що прогнозується. Довгостроковий прогноз меншою мірою залежить від початкового стану атмосфери і включає в себе елементи моделювання клімату. Збільшення завчасності прогнозів веде до необхідності розширення області інтегрування аж до охоплення всієї поверхні земної кулі, що висуває на перший план чинник часу виконання розрахунків і потужності комп'ютерів.

В атмосфері є процеси та явища різних типів. Одні процеси розвиваються на обмеженій території на протязі невеликого відрізка часу, інші – на широкому просторі довгий час. Найбільш розповсюдженою нині є шкала масштабів атмосферних процесів та об'єктів, згідно якої їх можна умовно поділити на три типи [13]: мікомасштабні (до 200 м), мезомасштабні (від 200 м до 2000 км) та макромасштабні (понад 2000 км). Зрозуміло, що формування процесів та явищ в атмосфері в кожній конкретній області виникає під впливом процесів всіх перерахованих вище масштабів. Але роль кожного процесу в різних умовах буде неоднаковою. Так, формування загальної циркуляції на великих просторах виникає головним чином під впливом макромасштабних процесів. Мезопроцеси при цьому вносять свій певний, але менш важливий внесок.

Раніше оперативні служби погоди розвинених країн світу концентрували свою увагу на моделях прогнозу синоптичних процесів на півкулі, незважаючи на низку результатів, згідно яких, успішність таких прогнозів зберігалася майже на постійному рівні. Найбільш досконалі обчислювальні ресурси, потрібні для реалізації оперативних моделей, зосереджувались у спеціально створених національних метеорологічних центрах. Для дослідження і прогнозу макромасштабних процесів використовуються макромасштабні (глобальні, півкульні) моделі. Такі моделі розроблені в усіх існуючих світових (Вашингтон, Мельбурн, Москва), а також у найбільших національних та регіональних метеорологічних центрах. Для реалізації макромасштабних моделей використовуються потужні комп'ютерні системи, що входять до списку top-500 [14].

Регіональні та локальні моделі, що описують мезомасштабні та мікомасштабні метеорологічні процеси та явища, визначають формування погоди в окремих регіонах та пунктах, призначені для створення деталізованих як в просторі, так і в часі прогнозів погоди. Останніми роками у більшості національних, регіональних і світових гідрометеорологічних службах створюються чи вже створені регіональні моделі і пакети програм для розв'язання задач чисельного прогнозу як окремих гідрометеорологічних явищ, так і погоди загалом. Ці моделі призначені для описання локальних процесів на відносно невеликих територіях і в майбутньому дадуть більший ефект, ніж традиційний прогноз погоди. Однак при розробці і впровадженні в оперативну практику регіональних моделей виникають надто важкі технологічні проблеми. Задача їх подолання сама стала задачею науки. Так, основними чинниками, що визначають поведінку багатьох мезомасштабних явищ в атмосфері, є тривимірна турбулентність і неоднорідність підстильної поверхні. Висновки, що випливають із теоретичних досліджень турбулентності [7,11,12] і аналізу вимірних спектрів, зводяться до того, що неминучі помилки і невизначеність у початкових полях дрібномасштабних рухів передаються рухам великих масштабів і досягають мезомасштабних збурень раніше від синоптичних, роблячи тим самим перші менш передбачуваними. Неоднорідність підстильної поверхні, у тому числі рельєф, альbedo, теплоємність і вміст вологості у ґрунті, генерують такі мезомасштабні явища, як гірські хвилі, морські бризи, конвекцію і зумовлюють їх подальшу поведінку. Зменшення неточностей, що модулюються наведеними чинниками, в мезомасштабних полях, які прогноуються внаслідок збільшення густоти розрахункової сітки, можливо реалізувати лише на дуже потужних комп'ютерах.

Прогнозування напрямків переміщення добре розвинутих вихрових утворень (наприклад, циклонів) та їх інтенсивності у гірських районах пов'язано з певними труднощами. У спокійному навколишньому середовищі циклон повільно зміщується під впливом адвекції планетарного вихору швидкості і, природно, неоднорідних термічних умов на підстильній поверхні. Коли ж циклон переміщується у район пасма високих крутих гір або починає взаємодіяти з іншими вихровими утвореннями, його траєкторію і швидкість прямування часто дуже важко визначити. Дослідження таких ситуацій за допомогою чисельних моделей набуває великого значення для актуальних практичних задач прогнозування небезпечних і стихійних метеорологічних явищ (тривалі інтенсивні опади, шквали, смерчі і, як результат, паводки, селі, сніжні лавини). Ключовою проблемою тут є надто високий ступінь деталізації моделі та надійне описання за її допомогою полів тих метеорологічних величин, що, у першу чергу, пов'язані з параметризацією неадіабатичних процесів.

Однією з головних особливостей карт погоди середніх широт є наявність атмосферних фронтів – поверхонь розриву, що розділяють повітряні маси з різними густинами, тобто зон більш різких градієнтів температури, тиску, вологості і складових швидкості вітру. Атмосферні фронти мають хвильовий характер. Горизонтальний масштаб фронтальних хвиль є на декілька порядків більшим за мезомасштабні особливості циркуляції у площині, що перпендикулярна до головної течії. Максимуми горизонтального градієнта температури і вертикальної швидкості мають місце поблизу поверхні Землі. Синхронна тривимірна структура мезомасштабних полів температури і швидкості вітру, що утворюється, обумовлює процеси, які генерують, підтримують і, нарешті, руйнують різкі градієнти температури на фронтальних поверхнях. Отже, моделювання цих мезомасштабних тонких і складних фізичних процесів фронтогенезу стає неможливим через втрату точності при чисельному інтегруванні рівнянь повної моделі атмосферної циркуляції з кроком сітки, що перевищує 10 км по горизонталі. Моделювання взаємодії конвективних систем із фронтальною циркуляцією і впливу планетарного прилежого шару є вкрай утрудненим. Питання про те, до якого ступеня треба подрібнювати скінчено-різницьову сітку, аби домогтися підвищення реалістичності математичної моделі, залишається досі не вирішеним, що вказує на все ще недостатнє розуміння процесів, що визначають динаміку фронтів.

Цілком зрозумілі ті труднощі, які виникають при моделюванні метеорологічних процесів різних масштабів з єдиних позицій. Наявність зв'язків між циркуляцією атмосфери та закономірностями обміну імпульсом, теплом

і вологою у приземному прилежовому шарі атмосфери безпосередньо проявляється при розв'язанні відповідних задач.

Основні математичні співвідношення, які описують мезомасштабні циркуляції, що генеруються термічними і орографічними чинниками, розроблені й апробовані. Методи динамічного прогнозування погоди й окремих метеорологічних явищ дуже ефективні й досить точні, щоб виправдати застосування їх у практиці, незважаючи на перелічені вище проблеми і хиби, які властиві існуючим математичним моделям циркуляції атмосфери і чисельним методам їх реалізації. Це підтверджує багаторічне їх використання в оперативній прогностичній роботі багатьох гідрометеорологічних центрів. Поліпшення потребують дискретизація рівнянь, методи їх чисельного розв'язання.

Підвищення завчасності й справджуваності прогнозів метеорологічних величин, небезпечних і стихійних явищ відкривають виняткові можливості і перспективи математичного моделювання і ще більшого його практичного застосування в задачах попередження і зменшення негативних наслідків від надзвичайних ситуацій природного і техногенного характеру. Математичне моделювання на сучасному етапі наукових досягнень і рівні технічного оснащення (чисельний експеримент, комп'ютерний сценарій та гідродинамічний прогноз) є однією з важливих і ефективних систем інформаційного забезпечення, експертного аналізу, обґрунтування організаційних рішень і підвищення наукового рівня керування. Розробка і впровадження безперервних прогнозувальних систем наразі стають все більш актуальними у зв'язку зі впровадженням автоматизованих систем оцінки ризику, прийняття рішень і керування.

Інша можливість математичного моделювання, яка ще у повній мірі не проявилася, – це його роль в експерименті, тобто наборі прогнозних сценаріїв, що сприяють підвищенню обґрунтованості, об'єктивності та ефективності розроблюваних на їх основі розв'язань. Якщо дослідження в інших науках звичайно реалізуються лабораторними експериментами, то велика кількість метеорологічних задач, що дотепер вважалися нерозв'язними, можуть бути розв'язаними на основі чисельних експериментів. Зазвичай, математичне моделювання проводиться в три етапи [12]:

- 1) побудова математичної моделі, що описує процеси в атмосфері;
- 2) вибір методу реалізації моделі;
- 3) програмування обчислювального алгоритму на ЕОМ і проведення чисельних розрахунків.

Щодо першого етапу, на сьогодні існує цілком задовільна математична теорія руху рідини взагалі й атмосферних рухів, зокрема [1–12]. Ця теорія ґрунтується на фізичних законах збереження кількості руху, маси й енергії. Математично вона виражається рівняннями руху Нав'є-Стокса, що зв'язують прискорення в певному напрямку з компонентами об'ємних і поверхневих сил, які діють у цьому напрямку, рівнянням збереження маси, термодинамічним рівнянням енергії і, нарешті, рівнянням стану Бойля-Шарля [11–12]. Ці фундаментальні закони, залишаючись істотно незмінними, відіграють ключову роль у розумінні математичних моделей прогнозування погоди.

Реалізація прогностичних моделей відбувається за допомогою чисельних методів. Це обумовлено складністю математичних моделей, що описують циркуляційні процеси в атмосфері. Рівняння, з яких складаються такі моделі, є переважно нелінійними трьохвимірними рівняннями другого порядку з малим параметром при похідних другого порядку. Завдяки наявності малого параметра при старших похідних ці рівняння можуть змінювати свій тип (еліптичний, гіперболічний або параболічний) залежно від режиму руху в атмосфері, що моделюється. Ресурси сучасних навіть дуже потужних комп'ютерних систем поки що не дають змоги здійснювати просторову дискретизацію моделі атмосфери з високим розділенням. А просторова дискретизація з грубим розділенням веде до того, що багато мезомасштабних метеорологічних процесів не описуються цією моделлю, оскільки у такому разі опиняються у класі процесів підсіткового масштабу.

Існують два методи описання механізмів взаємодії процесів різних масштабів. Перший ґрунтується на дискретизації мезомасштабної задачі на нерівномірній сітці з використанням необхідної інтерполяції метеорологічних параметрів, які задані на грубій сітці, у вузли дрібної сітки. Складність цього алгоритму очевидна. Другий метод полягає у розбиванні області розв'язання загальної задачі на низку підобластей, перетворенням у кожній підобласті постановки вихідної задачі з урахуванням описання лише основних процесів та зшиванням розв'язання на межах. Цей напрям нам видається більш економічним, тому конструкція основних алгоритмів, поданих нижче, будується з урахуванням можливості використання так званого методу «однобічного впливу» [12].

Значною проблемою при прогнозуванні погоди на основі реалізації математичних моделей є обмеження на час розв'язання, оскільки прогноз погоди потрібно отримувати завчасно. Навіть сучасні потужні обчислювальні машини не дозволяють використовувати дрібну просторову сітку в області розв'язання задачі і часову дискретизацію та ітераційні методи, що дало б можливість підвищити точність прогнозу. Особливо актуальною дана проблема видається для України, тому що питання зменшення затрат машинного часу при розв'язанні задач динамічної метеорології стоїть особливо гостро в зв'язку з відсутністю в національній гідрометслужбі потужної обчислювальної техніки, яку використовують у світових і національних метеорологічних центрах розвинутих держав.

Комп'ютерна реалізація чисельних методів для таких складних задач, як моделювання метеорологічних процесів, наразі вже неможлива без використання методів високопродуктивних обчислень. І справа не тільки в тому, що прогнозування погоди завжди пов'язане з жорсткими часовими обмеженнями, подолання яких потребує застосування потужних мультипроцесорних систем. Мережеві технології та засоби Інтернет,

операційні системи і прикладне програмне забезпечення в сучасних умовах так чи інакше ґрунтуються на концепціях паралельних і розподілених обчислень, використання яких є важливим для досягнення високих показників продуктивності мультипроцесорної обробки і паралельного програмування.

Задача моделювання метеорологічних процесів передбачає для свого розв'язання розщеплення рівнянь за координатами [10, 15] і так званий природний паралелізм виконання операцій [12], тобто такий паралелізм виконання операцій, що допускається специфікацією самої задачі, а не впливає з алгоритму її розв'язання. Одним з перспективних напрямів у розробці і дослідженні систем паралельних і розподілених обчислень нині є побудова програмних абстракцій у вигляді алгеброалгоритмічних мов і моделей, що ставить своєю метою розвиток архітектурно- і мовно-незалежних засобів програмування для мультипроцесорних обчислювальних систем і мереж. Застосування і розвиток алгеброалгоритмічного підходу до побудови структурованих абстракцій у моделях паралельних програм має на меті забезпечення властивостей високої продуктивності паралельних обчислень.

Подальший прогрес у покращенні показників якості створення паралельних систем метеорологічного прогнозування пов'язаний не тільки з підвищенням продуктивності їх роботи, а й зі зменшенням споживаної енергії на одиницю обчислювальної потужності. Досягнення таких цілей лежить на шляху до використання неоднорідних архітектур паралельних обчислювальних систем, здатних до спеціалізації обчислень для конкретних класів задач з метою одержання найвищих показників продуктивності при найменших ресурсних витратах. В останні роки у цьому відношенні з найкращого боку себе проявили, наприклад, відеографічні прискорювачі (GPU) – спеціалізовані недорогі мультипроцесорні системи, спроектовані спочатку як додаткові до центральних процесорів засоби обробки відеоінформації і використані потім у якості мультипроцесорних платформ загального призначення (GPGPU). Проте, як відомо, програмувати неоднорідні архітектури значно складніше, ніж однорідні, тому постає питання щодо розробки спеціальних засобів автоматизації розробки програмного забезпечення, що дозволяли б найбільш ефективно генерувати найпродуктивніший код для таких систем. Відмітимо, що поняття неоднорідних архітектур паралельних систем виникає також в зв'язку з побудовою глобальних інформаційно-обчислювальних інфраструктур, таких як GRID-системи і “хмарні” обчислення (cloud computing), що використовують низку взаємно пов'язаних ідей, включаючи поняття утилітарних (загальнодоступних модульних) обчислень, сервісів, GRID-технологій, віртуальних ресурсів та ін., спрямованих на граничне абстрагування ресурсів та спрощення доступу до них з боку користувача.

У даній статті запропоновано підхід до вирішення проблем чисельного прогнозування погоди, що поєднує використання адекватних фізичних моделей атмосферних процесів з ефективними обчислювальними схемами і методами програмування високопродуктивних обчислень на мультипроцесорних системах, а також методами машинного навчання, що дає змогу досягати належного ступеня точності, повноти і своєчасності метеорологічних прогнозів.

Підхід нейронних мереж

Як зазначалося вище, циркуляційні процеси в атмосфері, що мають вплив на метеорологічні величини, описуються складними математичними моделями. Для отримання точного прогнозу необхідно здійснити просторову дискретизацію моделі атмосфери з високим розділенням, оскільки просторова дискретизація з грубим розділенням веде до того, що багато мікомасштабних метеорологічних процесів у такому разі опиняються у класі процесів підсіткового масштабу. Проте просторова дискретизація з високим розділенням потребує залучення ресурсів, що наразі не можуть бути надані навіть дуже потужними комп'ютерними системами.

Звичайно, обчислювальна потужність комп'ютерних систем зростає все більше з розвитком технологій і в майбутньому може стати достатньою для того, щоб здійснювати дискретизацію з високим розділенням і, як наслідок, отримувати точні прогнози. Проте створення і обслуговування подібної комп'ютерної системи матиме високу вартість, і тому її використання може бути нерентабельним. Саме тому актуальним є залучення інших технік і методів для отримання прогнозів більш високої точності.

Однією з таких технік уточнення прогнозу чисельної моделі може бути використання машинного навчання, а саме нейронних мереж. Сьогодні спектр використання нейронних мереж є широким, що спричинено розвитком нових алгоритмів навчання і збільшенням обчислювальної потужності комп'ютерних систем. Причиною популярності нейронних мереж є той факт, що вони дозволяють знаходити загальний розподіл, зберігаючи інформацію про локальні деталі [16].

Дана стаття описує архітектуру нейронної мережі для задачі уточнення чисельного прогнозу температури. Запропонована мережа прогнозує помилки прогнозів чисельної моделі з метою подальшого коригування власне прогнозу.

Опис даних. Здатність нейронної мережі розв'язувати ту чи іншу задачу суттєво залежить від даних, які використовуються для її навчання, їх репрезентативності і цілісності.

Дані, які використовувались для навчання, валідації і тестування нейронної мережі, склалися з чотирьохелементних кортежів і містили наступну інформацію:

- дата,
- час по Гринвічу,

– прогнозоване значення температури (Fest), завчасістю в одну добу від моменту ініціалізації чисельної регіональної моделі,

– спостережуване значення температури (Obs) по метеостанції Київ.

В якості регіональної чисельної моделі прогнозу погоди було використано модель COSMO, однойменного європейського консорціуму (COSMO розшифровується як Консорціум з моделювання на малих масштабах), яка в Українському гідрометеорологічному інституті ДСНС України та НАН України використовується для наукових та прикладних задач, починаючи із липня 2011 р. [17,18]. COSMO є негідростатичною моделлю, яка здатна ефективно відтворювати широкий спектр атмосферних процесів в масштабі мезо- β та мезо- γ . В основу динамічного ядра моделі покладено рівняння термо- та гідродинаміки, що описують потік у вологій атмосфері. Різноманітні фізичні процеси враховуються схемами параметризації [17]. Розрахункова область (Рис.1) характеризується наступним чином: кількість вузлів із заходу на схід – 209; кількість вузлів із півдня на північ – 101; кількість рівнів по вертикалі – 50; крок ~ 14 км.



Рис. 1. Розрахункова область чисельної регіональної моделі прогнозу погоди COSMO

Дані спостережень проводилися кожні три години, а саме в 00:00, 03:00, 06:00, 09:00, 12:00, 15:00, 18:00 і 21:00 по Гринвічу. Для цих моментів часу обчислювався і прогноз регіональної моделі. Таким чином, для кожної дати маємо по вісім кортежів. Дані охоплюють проміжок часу від 01.07.2012 до 01.11.2013, тобто маємо 489 днів і, відповідно, 3912 кортежів.

Для тренування нейронної мережі уся сукупність даних була розбита на три класи: тренувальні (період з 01.07.2012 по 30.06.2013; 365 днів), тестувальні (період з 01.07.2013 по 24.10.2013; 116 днів) і валідаційні (з 25.10.2013 по 01.11.2013; 8 днів). Співвідношення тренувальних і тестувальних даних приблизно 3:1.

Вхідні і вихідні вектори. Як вже зазначалося, основна ідея – натренувати нейронну мережу на помилках прогнозів регіональної моделі, тобто на значеннях $Error = F_{cst} - Obs$. Тоді маючи прогноз регіональної моделі і прогнозоване значення помилки, уточнене значення прогнозу F_{cstNN} обчислюється як $F_{cst} - Error$.

Експерименти проводилися з різними передісторіями помилок, а саме з помилками за 1, 2, 3, 5 і 7 діб. Таким чином, розмірність вхідного вектора дорівнювала, відповідно, 1×8 , 2×8 , 3×8 , 5×8 і 7×8 для прогнозування помилок прогнозів регіональної моделі різної завчасності (8 вимірювань на добу). Тобто кожен вхідний вектор містив повну інформацію про помилку регіональної моделі за відповідний період.

Розмірність вихідного вектора завжди дорівнювала 1×8 . Він містив повну інформацію для уточнення прогнозу регіональної моделі для однієї доби.

Архітектура. Оскільки дані мають природу часового ряду, першим шаром нейронної мережі було вирішено взяти рекурентний, оскільки рекурентні мережі створені саме для роботи з даними такого роду. З-поміж наявних видів рекурентних архітектур була обрана архітектура вентильного рекурентного вузла GRU [19],

оскільки вона містить небагато параметрів (порівняно з популярною архітектурою з довгою короткочасною пам'яттю LSTM [20]), а тому навчається швидше. При цьому вона дозволяє уникнути основних проблем, що виникають при аналізі довгострокових залежностей, а саме проблеми ефекту зникнення градієнтів (vanishing gradient effect) і проблеми безпосередньо пов'язаного з ним ефекту вибуху градієнтів (exploding gradients effect). Вузол GRU працює наступним чином. В кожний момент часу йому на вхід подаються вхідний вектор x_t і вихідний вектор з попереднього моменту часу h_{t-1} . Вихідний вектор h_t обчислюється як лінійна інтерполяція між h_{t-1} і поточним кандидатом \tilde{h}_t :

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t,$$

де z_t – вектор вузла уточнення, що містить значення, які визначають, які частини першого і другого векторів впливатимуть на поточне значення, а \odot позначає добуток Адамара. Цей вектор обчислюється наступним чином:

$$z_t = \sigma_z(W_z x_t + U_z h_{t-1} + b_z).$$

Поточний кандидат \tilde{h}_t обчислюється подібно до стандартної RNN:

$$\tilde{h}_t = \sigma_h(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h),$$

де r_t – вектор вузла скидання, що обчислюється подібно до вектора вузла уточнення:

$$r_t = \sigma_r(W_r x_t + U_r h_{t-1} + b_r).$$

$W_z, U_z, W_h, U_h, W_r, U_r$ позначають матриці параметрів, а b_z, b_h, b_r позначають вектори параметрів. Вектори вузлів уточнення і скидання активуються сигмоїдною функцією, а вектор поточного кандидату – гіперболічним тангенсом.

В рамках поточної архітектури, для прогнозів усіх завчасностей першим шаром нейронної мережі було взято GRU шар, що складається з 16 вузлів, оскільки менша і більша кількість давали гірші результати.

Після GRU шару було додано два прихованих повнозв'язних шари. Для прогнозів різної завчасності вони мали різну кількість вузлів. Для прогнозів з передісторіями 1–2 доби вони мали, відповідно, 8 і 16 вузлів, для прогнозів з передісторіями 3–5 діб – 32 і 16 вузлів, а для прогнозу з тижневою передісторією – 8 і 32 вузли, відповідно.

Останній шар запропонованої архітектури був також повнозв'язним і мав 8 вузлів, оскільки результуючий вектор, що був отриманий при проходженні всіх шарів, повинен мати розмірність вихідного вектора.

Усі експерименти проводилися з використанням відкритого нейромережевого інтерфейсу Keras і відкритої програмної бібліотеки для машинного навчання TensorFlow [21]. Програмний код був написаний мовою Python.

Експерименти

Як вже зазначалося, експерименти були проведені з різними передісторіями помилок регіональної моделі, а саме з помилками за 1, 2, 3, 5 і 7 діб. Навчалися моделі різну кількість епох: 15 (для завчасностей у 1 і 2 доби) і 6 (для завчасностей у 3, 5 і 7 діб). Більша кількість (ніж запропоновано у даній роботі) епох призводила до так званого перенавчання мережі і, як наслідок, давала більші значення середніх квадратів помилок (MSE) на валідаційних і тестових даних. Графіки середніх квадратів помилок під час тренування і валідації наведені на Рис. 2.

Пунктирною лінією позначено помилку на тренувальних даних, а суцільною лінією – на валідаційних даних: а) – передісторія в 1 добу; б) – передісторія у 2 доби; в) – передісторія у 3 доби; г) – передісторія у 5 діб; д) – передісторія у 7 діб дозволяє оцінити ефективність застосування власне архітектури нейронної мережі для різних часових проміжків. Із результатів експериментів, показаних на Рис. 2 та у таблиці, випливає, що найкращою є використання передісторії в одну добу, тобто синхронно із чисельною регіональною моделлю погоди. Із збільшенням передісторій відсоток випадків, в яких прогноз покращено, поступово зменшується. Ефект збільшення відсотку покращених прогнозів для передісторій в 5 та 7 діб потребує подальших досліджень.

Таблиця. Відсоток випадків, в яких прогноз покращено. Дані наводяться для різних передісторій помилок

Передісторія, доба	Відсоток покращених прогнозів, %
1	52.05
2	50.98
3	50.44
5	51.67
7	52.04

Необхідно відзначити, що для тренувального періоду найменші значення MSE мали місце при максимальному запропонованому значенні епох для кожної із передісторій. Отримані значення MSE є цілком співставними із результатами застосування крігінг-інтерполяції [22]. Щоправда, використання нейронних мереж видається більш перспективним підходом, оскільки дає можливість коригувати помилки викликані не лише унаслідок інтерполяції прогнозу із вузлів регулярної (модель) сітки до точок нерегулярної сітки (метеорологічні станції), так і особливостями підстильної поверхні та фізичних процесів, які не враховуються або частково враховуються у моделі прогнозу погоди.

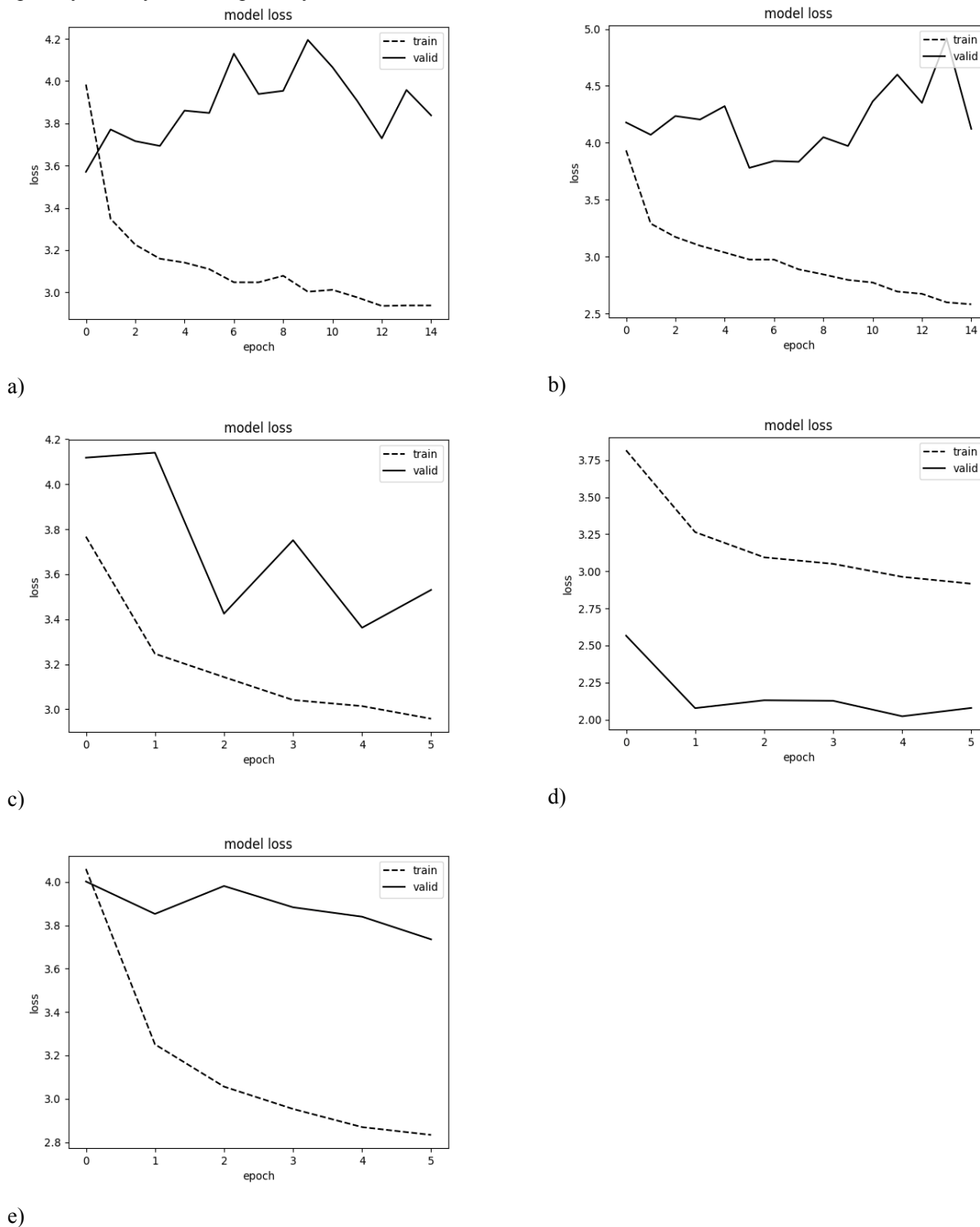


Рис. 2. Середній квадрат помилки для передісторії різної тривалості

Висновки

На прикладі прогнозів приземної температури повітря моделі COSMO для м. Київ та відповідних їм даних фактичних спостережень запропоновано архітектуру нейронної мережі, яка дозволяє прогнозувати помилки власне прогнозів чисельної регіональної моделі, з метою подальшого коригування прогнозу неперервних метеорологічних величин.

Проведено експерименти з різними передісторіями помилок регіональної моделі, а саме з помилками за 1, 2, 3, 5 і 7 діб. Визначено кількість епох, після збільшення якої мало місце так зване перенавчання мережі: 15 (для завчасностей у 1 і 2 доби) і 6 (для завчасностей у 3, 5 і 7 діб).

Показано, що запропонована архітектура дає можливість досягти покращення приземної температури повітря приблизно у 50 % випадків прогнозів.

Література

1. Андерсон Д., Таннехилл Дж., Плетчер Р. Вычислительная гидромеханика и теплообмен. / Под ред. Г. В. Подвиза. М.: Мир, 1990. Т. 2. 726 с.
2. Белов П. Н., Борисенков Е.П., Панин Б.Д. Численные методы прогноза погоды. Л.: Гидрометеиздат, 1989. 376 с.
3. Володин Е. М. Математическое моделирование общей циркуляции атмосферы. М.: ИВМ РАН, 2007. 89 с.
4. Гилл А. Динамика атмосферы и океана. В 2-х т. М.: Мир, 1986. 416 с.
5. Гладкий А. В., Скопеевский В. В. Методи числового моделювання екологічних процесів. К.: ІВЦ „Політехніка”, 2005. 152 с.
6. Динамика погоды. / Под ред. С. Манабе. Л.: Гидрометеиздат, 1988. 420 с.
7. Дымников В. П. Устойчивость и предсказуемость крупномасштабных атмосферных процессов. М.: ИВМ РАН, 2007. 283 с.
8. Ківганов А. Ф., Хоменко Г. В., Хохлов В. М., Бондаренко В.М. Гідродинамічні методи прогнозу погоди і сіткові методи їх реалізації. Одеса: Одеський державний екологічний університет, 2002. 179 с.
9. Козаков О. Л. Динамічна метеорологія. Одеса: ОДЕКУ, 2003. 148 с.
10. Марчук Г.И. Численные методы в прогнозе погоды. Л.: Гидрометеиздат, 1967. 353 с.
11. Прусов В.А., Сніжко С.І. Математичне моделювання атмосферних процесів. К.: Ніка-Центр. 2005. 496 с.
12. Прусов В.А., Дорошенко А.Ю. Моделювання природних і техногенних процесів в атмосфері. К: Наукова Думка. 2006. 542 с.
13. Orlandi I., 1975. A Rational Subdivision of Scales for Atmospheric Processes. *Bulletin of the American Meteorological Society*, 56(5). P. 527–530.
14. www.top500.org
15. Прусов В.А., Дорошенко А.Е., Черныш Р.И., Гук Л.Н. Эффективная разностная схема численного решения задачи конвективной диффузии. *Кибернетика и системный анализ*, 2007. 3, P. 64–74.
16. Frnda J., Durica M., Nedoma J., Zabka S., Martinek R. and Kostelansky M. A Weather Forecast Model Accuracy Analysis and ECMWF Enhancement Proposal by Neural Network. *Sensors*, 2019. 19(23). P. 5144.
17. Doms G.A Description of the Nonhydrostatic Regional COSMO-Model. Part I: Dynamics and Numerics. Offenbach: Deutscher Wetterdienst. 2011. 153 p.
18. Shpyg V. et al. The application of regional NWP models to operational weather forecasting in Ukraine. CAS Technical Conference (TECO) on “Responding to the Environmental Stressors of the 21st Century”. 2013.
URL: <http://www.wmo.int/pages/prog/arep/cas/documents/Ukraine-NWPModels.pdf> (дата звернення: 27.02.2020).
19. Cho K., Van Merriënboer B., Gulcehre C., Bahdanau D., Bougares F., Schwenk H. and Bengio Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv*: 2014. 1406.1078.
20. Hochreiter S. and Schmidhuber J. Long short-term memory. *Neural computation*. 1997. 9(8). P. 1735–1780.
21. www.tensorflow.org
22. Katsalova L.M. and Shpyg V.M., The choice of optimal lag for Kriging interpolation of NWP model forecast. *Meteorology, Hydrology and Water Management*. 2016. 4(2). P. 23–28.

References

1. Anderson D. & Tannehill J. & Pletcher R. (1990). Computational fluid mechanics and heat transfer. Ed. by G. V. Podviz. Moscow: Mir. 726 p. (in Russian).
2. Belov P.N. & Borisenkov E.P. & Panin B.D. (1989) Methods of Numerical Weather Prediction. Leningrad:Gidrometeoizdat. 376 p. (in Russian).
3. Volodin E.M. (2007). Mathematical modeling of the general circulation of the atmosphere. Moscow: Institute of Computational Mathematics, Russian Academy of Sciences. 89 p. (in Russian).
4. Gill Adrian E. (1986) Atmosphere-Ocean Dynamics. Vol. 1,2. Moscow: Mir. 416 p. (in Russian).
5. Gladkyi A.V. & Skopetsky V.V. (2005) Methods of numerical modeling of ecological processes. Kyiv: Politekhnik. 152 p. (in Ukrainian).
6. (1988). Weather Dynamics. Ed. by S. Manabe. – Leningrad.:Gidrometeoizdat. 420 p. (in Russian).
7. Dymnikov V.P. (2007). Stability and predictability of large-scale atmospheric processes. Moscow: Institute of Computational Mathematics, Russian Academy of Sciences. 283 p. (in Russian).
8. kivganov A.F. et al. (2002). Hydrodynamic weather forecasting methods and grid methods for their implementation. Odessa: Odessa State Environmental University. 179 p. (in Ukrainian).
9. Kozakov O.L. (2003). Dynamical Meteorology. Odessa: Odessa State Environmental University. 148 p. (in Ukrainian).
10. Marchuk G.I. (1967). Numerical Methods in Weather Prediction. Leningrad.:Gidrometeoizdat. 353 p. (in Russian).
11. Prusov V.A. & Snizhko S.I. (2005). Mathematical modeling of atmospheric processes. Kyiv: Nika-Center. 496 p. (In Ukrainian).
12. Prusov V.A. & Doroshenko A.Yu. (2006). Modeling of natural and technogenic processes in the atmosphere. Kyiv: Naukova dumka. 542 p. (in Ukrainian).
13. Orlandi I. (1975). A Rational Subdivision of Scales for Atmospheric Processes. Bulletin of the American Meteorological Society. 56(5). P. 527–530.
14. www.top500.org
15. Prusov V.A. et. al. (2007). Effective difference scheme for the numerical solution of the convective diffusion problem. Cybernetics and Systems Analysis. 3. P. 64–74.

16. Frnda J. et al. (2019). A Weather Forecast Model Accuracy Analysis and ECMWF Enhancement Proposal by Neural Network. *Sensors*. 19(23). P. 5144.
17. Doms G. (2011). A Description of the Nonhydrostatic Regional COSMO-Model. Part I: Dynamics and Numerics. Offenbach: Deutscher Wetterdienst. 153 p.
18. Shpyg V. et al. (2013). The application of regional NWP models to operational weather forecasting in Ukraine. CAS Technical Conference (TECO) on "Responding to the Environmental Stressors of the 21st Century". Available from: <http://www.wmo.int/pages/prog/arep/cas/documents/Ukraine-NWPModels.pdf> [Accessed 27/02/2020].
19. Cho K. et al. (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.
20. Hochreiter S. & Schmidhuber J. (1997). Long short-term memory. *Neural computation*, 9(8), P. 1735–1780.
21. www.tensorflow.org
22. Katsalova L.M. & Shpyg V.M. (2016). The choice of optimal lag for Kriging interpolation of NWP model forecast. *Meteorology, Hydrology and Water Management*. 4(2). P. 23–28.

Одержано 11.03.2020

Про авторів:

Дорошенко Анатолій Юхимович,

доктор фізико-математичних наук, професор, завідувач відділу,
професор кафедри автоматизації та управління в технічних системах КПІ імені Ігоря Сікорського.
Кількість наукових публікацій в українських виданнях – понад 180.
Кількість наукових публікацій в зарубіжних виданнях – понад 60.
<http://orcid.org/0000-0002-8435-1451>,

Шпиг Віталій Михайлович,

кандидат географічних наук,
завідувач відділу фізики атмосфери
Українського гідрометеорологічного інституту ДСНС України та НАН України,
експерт Комісії з атмосферних наук Всесвітньої метеорологічної організації.
Кількість наукових публікацій в українських виданнях – понад 50.
Кількість наукових публікацій в зарубіжних виданнях – понад 50.
<https://orcid.org/0000-0003-1055-7120>,

Кушніренко Роман Владиславович,
аспірант.

Кількість наукових публікацій в українських виданнях – 1.
<https://orcid.org/0000-0002-1990-8727>.

Місце роботи авторів:

Інститут програмних систем НАН України,
03187, м. Київ-187,
проспект Академіка Глушкова, 40.
Тел.: (38)(044) 526-60-33.
E-mail: doroshenkoanatoliy2@gmail.com,
romashka1996@gmail.com.

Український гідрометеорологічний інститут ДСНС України та НАН України,
03028, м. Київ, проспект Науки, 37.
Тел.: (38)(044) 525-86-30.
E-mail: vitold82@i.ua.