城市气象科技信息监测简报

（文献类）

2022 年 10 月（总第 16 期）

北京气象学会
北京城市气象研究院

2022 年 11 月 8 日

前 言

2021年7月，依托北京城市气象研究院行业信息监测与分析系统，北京气象学会与支撑单位北京城市气象研究院联合制作了《城市气象科技信息监测简报》，定期为会员提供科技信息服务。简报每月一期，每期分为文献和咨询两类。

目前，学会主要依据城市气象领域的关键词搜索信息，并形成每月简报。欢迎各位会员向我们提供相关科技领域信息及其搜索关键词，不断扩充简报的专业领域范围，丰富简报的内容，以更好地满足广大会员朋友们对快速更新科技信息的需求。

同时，热诚欢迎广大会员对简报的科技信息内容、展现形式、阅读体验和收获感悟等提出建议和点评。编辑团队将遴选出优秀建议和点评内容刊载在简报上。

2020年12月，学会换届成立了第21届理事会，现拥有52 家理事单位和 947 名会员。我们希望借助此简报，为广大会员朋友们提供高质高效的科技信息服务，同时在大家的支持和帮助下，我们共同将简报建设成为会员交流的友好平台。

编辑团队:

苏　晨 北京城市气象研究院办公室副主任

楚艳丽　北京气象学会副秘书长、北京城市气象研究院副研究员

苗世光　北京气象学会秘书长、北京城市气象研究院院长

郭佩芸 北京城市气象研究院专职副书记

胡永云 北京气象学会理事长、北京大学教授

刘　燕 北京气象学会副秘书长

目 录

❖报告概述 2

✧监测周期 2

✧监测主题 2

❖专题研究 2

✧城市边界层 2

1. Development of a regional feature selection-based machine learning system (RFSML v1.0) for air pollution forecasting over China 2
2. A new bootstrap technique to quantify uncertainty in estimates of ground surface temperature and ground heat flux histories from geothermal data 3
3. A tool for air pollution scenarios (TAPS v1.0) to enable global, long-term, and flexible study of climate and air quality policies 4
4. Recovery of sparse urban greenhouse gas emissions 4

✧城市气象精细预报 5

1. Data assimilation for the Model for Prediction Across Scales – Atmosphere with the Joint Effort for Data assimilation Integration (JEDI-MPAS 1.0.0): EnVar implementation and evaluation 5
2. A preliminary evaluation of FY-4A visible radiance data assimilation by the WRF (ARW v4.1.1)/DART (Manhattan release v9.8.0)-RTTOV (v12.3) system for a tropical storm case 6

✧城市气候与生态 7

1. A modeling framework to understand historical and projected ocean climate change in large coupled ensembles 7

✧人工智能气象 8

1. Repeatable high-resolution statistical downscaling through deep learning 8
2. Development of a regional feature selection-based machine learning system (RFSML v1.0) for air pollution forecasting over China 9

❖报告概述

✧监测周期

2022 年 10 月 01 日 – 2022 年 10 月 31 日

✧监测主题

专题研究——城市边界层、城市气象精细预报、城市气候与生态、人工智能气象；

❖专题研究

✧城市边界层

本监测周期内，文献信息共更新 4 条：

1. Development of a regional feature selection-based machine learning system (RFSML v1.0) for air pollution forecasting over China

摘要：With the explosive growth of atmospheric data, machine learning models have achieved great success in air pollution forecasting because of their higher computational efficiency than the traditional chemical transport models. However, in previous studies, new prediction algorithms have only been tested at stations or in a small region; a large-scale air quality forecasting model remains lacking to date. Huge dimensionality also means that redundant input data may lead to increased complexity and therefore the over-fitting of machine learning models. Feature selection is a key topic in machine learning development, but it has not yet been explored in atmosphere-related applications. In this work, a regional feature selection-based machine learning (RFSML) system was developed, which is capable of predicting air quality in the short term with high accuracy at the national scale. Ensemble-Shapley additive global importance analysis is combined with the RFSML system to extract significant regional features and eliminate redundant variables at an affordable computational expense. The significance of the regional features is also explained physically. Compared with a standard machine learning system fed with relative features, the RFSML system driven by the selected key features results in superior interpretability, less training time, and more accurate predictions. This study also provides insights into the difference in interpretability among machine learning models (i.e., random forest, gradient boosting, and multi-layer perceptron models).

来源：GMD

发布时间：2022-10-24

数据类型：期刊

https://gmd.copernicus.org/articles/15/7791/2022/

2. A new bootstrap technique to quantify uncertainty in estimates of ground surface temperature and ground heat flux histories from geothermal data

摘要：Estimates of the past thermal state of the land surface are crucial to assess the magnitude of current anthropogenic climate change as well as to assess the ability of Earth System Models (ESMs) to forecast the evolution of the climate near the ground, which is not included in standard meteorological records. Subsurface temperature reacts to long-term changes in surface energy balance – from decadal to millennial time scales – thus constituting an important record of the dynamics of the climate system that contributes, with low-frequency information, to proxy-based paleoclimatic reconstructions. Broadly used techniques to retrieve past temperature and heat flux histories from subsurface temperature profiles based on a singular value decomposition (SVD) algorithm were able to provide robust global estimates for the last millennium, but the approaches used to derive the corresponding 95 % confidence interval were wrong from a statistical point of view in addition to being difficult to interpret. To alleviate the lack of a meaningful framework for estimating uncertainties in past temperature and heat flux histories at regional and global scales, we combine a new bootstrapping sampling strategy with the broadly used SVD algorithm and assess its performance against the original SVD technique and another technique based on generating perturbed parameter ensembles of inversions. The new bootstrap approach is able to reproduce the prescribed surface temperature series used to derive an artificial profile. Bootstrap results are also in agreement with the global mean surface temperature history and the global mean heat flux history retrieved in previous studies. Furthermore, the new bootstrap technique provides a meaningful uncertainty range for the inversion of large sets of subsurface temperature profiles. We suggest the use of this new approach particularly for aggregating results from a number of individual profiles, and to this end, we release the programs used to derive all inversions in this study as a suite of codes labeled CIBOR v1: Codes for Inverting BORholes, version 1.

来源：GMD

发布时间：2022-10-28

数据类型：期刊

https://gmd.copernicus.org/articles/15/7913/2022/

3. A tool for air pollution scenarios (TAPS v1.0) to enable global, long-term, and flexible study of climate and air quality policies

摘要：Air pollution is a major sustainability challenge – and future anthropogenic precursor and greenhouse gas (GHG) emissions will greatly affect human well-being. While mitigating climate change can reduce air pollution both directly and indirectly, distinct policy levers can affect these two interconnected sustainability issues across a wide range of scenarios. We help to assess such issues by presenting a public Tool for Air Pollution Scenarios (TAPS) that can flexibly assess pollutant emissions from a variety of climate and air quality actions, through the tool's coupling with socioeconomic modeling of climate change mitigation. In this study, we develop and implement TAPS with three components: recent global and fuel-specific anthropogenic emissions inventories, scenarios of emitting activities to 2100 from the MIT Economic Projection and Policy Analysis (EPPA) model, and emissions intensity trends based on recent scenario data from the Greenhouse Gas–Air Pollution Interactions and Synergies (GAINS) model. An initial application shows that in scenarios with less climate and pollution policy ambition, near-term air quality improvements from existing policies are eclipsed by long-term emissions increases – particularly from industrial processes that combine sharp production growth with less stringent pollution controls in developing regions. Additional climate actions would substantially reduce air pollutant emissions related to fossil fuel (such as sulfur and nitrogen oxides), while further pollution controls would lead to larger reductions for ammonia and organic carbon (OC). Future applications of TAPS could explore diverse regional and global policies that affect these emissions, using pollutant emissions results to drive global atmospheric chemical transport models to study the scenarios' health impacts.

来源：GMD

发布时间：2022-10-21

数据类型：期刊

https://gmd.copernicus.org/articles/15/7767/2022/

4. Recovery of sparse urban greenhouse gas emissions

摘 要：To localize and quantify greenhouse gas emissions from cities, gas concentrations are typically measured at a small number of sites and then linked to emission fluxes using atmospheric transport models. Solving this inverse problem is challenging because the system of equations often has no unique solution and the solution can be sensitive to noise. A common top–down approach for solving this problem is Bayesian inversion with the assumption of a multivariate Gaussian distribution as the prior emission field. However, such an assumption has drawbacks when the assumed spatial emissions are incorrect or not Gaussian distributed. In our work, we investigate sparse reconstruction (SR), an alternative reconstruction method that can achieve reasonable estimations without using a prior emission field by making the assumption that the emission field is sparse. We show that this assumption is generally true for the cities we investigated and that the use of the discrete wavelet transform helps to make the urban emission field even more sparse. To evaluate the performance of SR, we created concentration data by applying an atmospheric forward transport model to CO2 emission inventories of several major European cities. We used SR to locate and quantify the emission sources by applying compressed sensing theory and compared the results to regularized least squares (LSs) methods. Our results show that SR requires fewer measurements than LS methods and that SR is better at localizing and quantifying unknown emitters.

来源：GMD

发布时间：2022-10-17

数据类型：期刊

https://gmd.copernicus.org/articles/15/7533/2022/

✧城市气象精细预报

本监测周期内，文献信息共更新 2 条：

1. Data assimilation for the Model for Prediction Across Scales – Atmosphere with the Joint Effort for Data assimilation Integration (JEDI-MPAS 1.0.0): EnVar implementation and evaluation

摘要：On 24 September 2021, JEDI-MPAS 1.0.0, a new data assimilation (DA) system for the Model Prediction Across Scales – Atmosphere (MPAS-A) built on the software framework of the Joint Effort for Data assimilation Integration (JEDI) was publicly released for community use. Operating directly on the native MPAS unstructured

mesh, JEDI-MPAS capabilities include three-dimensional variational (3DVar) and ensemble–variational (EnVar) schemes as well as the ensemble of DA (EDA) technique. On the observation side, one advanced feature in JEDI-MPAS is the full all-sky approach for satellite radiance DA with the introduction of hydrometeor analysis variables. This paper describes the formulation and implementation of EnVar for JEDI-MPAS. JEDI-MPAS 1.0.0 is evaluated with month-long cycling 3DEnVar experiments with a global 30–60 km dual-resolution configuration. The robustness and credible performance of JEDI-MPAS are demonstrated by establishing a benchmark non-radiance DA experiment, then incrementally adding microwave radiances from three sources: Advanced Microwave Sounding Unit-A (AMSU-A) temperature sounding channels in clear-sky scenes, AMSU-A window channels in all-sky scenes, and Microwave Humidity Sounder (MHS) water vapor channels in clear-sky scenes. JEDI-MPAS 3DEnVar behaves well with a substantial and significant positive impact obtained for almost all aspects of forecast verification when progressively adding more microwave radiance data. In particular, the day 5 forecast of the best-performing JEDI-MPAS experiment yields an anomaly correlation coefficient (ACC) of 0.8 for 500 hPa geopotential height, a gap of roughly a half day when compared to cold-start forecasts initialized from operational analyses of the National Centers for Environmental Prediction, whose ACC does not drop to 0.8 until a lead time of 5.5 d. This indicates JEDI-MPAS's great potential for both research and operations.

来源：GMD

发布时间：2022-10-26

数据类型：期刊

https://gmd.copernicus.org/articles/15/7859/2022/

2. A preliminary evaluation of FY-4A visible radiance data assimilation by the WRF (ARW v4.1.1)/DART (Manhattan release v9.8.0)-RTTOV (v12.3) system for a tropical storm case

摘要：Satellite visible radiance data that contain rich cloud and precipitation information are increasingly assimilated to improve the forecasts of numerical weather prediction models. This study evaluates the Data Assimilation Research Testbed (DART, Manhattan release v9.8.0), coupled with the Weather Research and Forecasting (WRF) model (ARW v4.1.1) and the Radiative Transfer for TOVS (RTTOV, v12.3) package, for assimilating the simulated visible imagery of the FY-4A geostationary satellite located over Asia in an Observing System Simulation Experiment (OSSE) framework. The OSSE was

performed for the tropical storm Higos that occurred in 2020 and contains multi-layer mixed-phase cloud and precipitation processes. The advantages and limitations of DART for assimilating FY-4A visible imagery were evaluated. Both single-observation experiments and cycled data assimilation (DA) experiments were performed to study the impact of different filter algorithms available in DART, variables being cycled, observation outlier thresholds, observation errors, and observation thinning.The results show that assimilating visible radiance data significantly improves the analysis of the cloud water path (CWP) and cloud coverage (CFC) from first-guess forecasts. The rank histogram filter (RHF) allows WRF to more accurately simulate CWP and CFC compared with the ensemble adjustment Kalman filter (EAKF) although it took roughly twice as long as the latter. By cycling both cloud and non-cloud variables, specifying large outlier threshold values, or setting smaller observation errors without thinning of observations, WRF achieved a better simulation of CWP and CFC. With model integration, DA of the visible radiance data also generated a slightly positive impact on non-cloud variables as they were adjusted through the model dynamics and physics related to cloud processes. In addition, the DA improved the representation of precipitation. However, the impact on the rain rate is limited by the inabilities of the DA to improve cloud vertical structures and cloud phases. A negative impact of the DA on cloud variables was found due to the nature of the non-linear forward operator and the non-Gaussian distribution of the prior. Future works should explore faster and more accurate forward operators suitable for assimilating FY-4A visible imagery, techniques to reduce the non-linear and non-Gaussian errors, and methods to correct the location errors which correspond to the clouds underestimated by the first guess.

来源：GMD

发布时间：2022-10-05

数据类型：期刊

https://gmd.copernicus.org/articles/15/7397/2022/

✧城市气候与生态

本监测周期内，文献信息共更新 1 条：

1. A modeling framework to understand historical and projected ocean climate change in large coupled ensembles

摘要：The ocean responds to climate change through modifications of heat, freshwater and momentum fluxes at its boundaries. Disentangling the specific role of each

of these contributors in shaping the changes of the thermohaline structure of the ocean is central for our process understanding of climate change and requires the design of specific numerical experiments. While it has been partly addressed by modeling studies using idealized CO2 forcings, the time evolution of these individual contributions during historical and projected climate change is however lacking. Here, we propose a novel modeling framework to isolate these contributions in coupled climate models for which large ensembles of historical and scenario simulations are available. The first step consists in reproducing a coupled pre-industrial control simulation with an ocean-only configuration, forced by prescribed fluxes at its interface, diagnosed from the coupled model. In a second step, we extract the external forcing perturbations from the historical+scenario ensemble of coupled simulations, and we add them to the prescribed fluxes of the ocean-only configuration. We then successfully replicate the ocean's response to historical and projected climate change in the coupled model during 1850–2100. In a third step, this full response is decomposed in sensitivity experiments in which the forcing perturbations are applied individually to the heat, freshwater and momentum fluxes. Passive tracers of temperature and salinity are implemented to discriminate the addition of heat and freshwater flux anomalies from the redistribution of pre-industrial heat and salt content in response to ocean circulation changes. Here, we first present this general framework and then apply it to the IPSL-CM6A-LR model and its ocean component NEMO3.6. This framework brings new opportunities to precisely explore the mechanisms driving historical and projected ocean changes within single climate models.

来源：GMD

发布时间：2022-10-21

数据类型：期刊

https://gmd.copernicus.org/articles/15/7683/2022/

✧人工智能气象

本监测周期内，文献信息共更新 2 条：

1. Repeatable high-resolution statistical downscaling through deep learning

摘要：One of the major obstacles for designing solutions against the imminent climate crisis is the scarcity of high spatio-temporal resolution model projections for variables such as precipitation. This kind of information is crucial for impact studies in

fields like hydrology, agronomy, ecology, and risk management. The currently highest spatial resolution datasets on a daily scale for projected conditions fail to represent complex local variability. We used deep-learning-based statistical downscaling methods to obtain daily 1 km resolution gridded data for precipitation in the Eastern Ore Mountains in Saxony, Germany. We built upon the well-established climate4R framework, while adding modifications to its base-code, and introducing skip connections-based deep learning architectures, such as U-Net and U-Net++. We also aimed to address the known general reproducibility issues by creating a containerized environment with multi-GPU (graphic processing unit) and TensorFlow's deterministic operations support. The perfect prognosis approach was applied using the ERA5 reanalysis and the ReKIS (Regional Climate Information System for Saxony, Saxony-Anhalt, and Thuringia) dataset. The results were validated with the robust VALUE framework. The introduced architectures show a clear performance improvement when compared to previous statistical downscaling benchmarks. The best performing architecture had a small increase in total number of parameters, in contrast with the benchmark, and a training time of less than 6 min with one NVIDIA A-100 GPU. Characteristics of the deep learning models configurations that promote their suitability for this specific task were identified, tested, and argued. Full model repeatability was achieved employing the same physical GPU, which is key to build trust in deep learning applications. The EURO-CORDEX dataset is meant to be coupled with the trained models to generate a high-resolution ensemble, which can serve as input to multi-purpose impact models.

来源：GMD

发布时间：2022-10-05

数据类型：期刊

https://gmd.copernicus.org/articles/15/7353/2022/

2. Development of a regional feature selection-based machine learning system (RFSML v1.0) for air pollution forecasting over China

摘要：With the explosive growth of atmospheric data, machine learning models have achieved great success in air pollution forecasting because of their higher computational efficiency than the traditional chemical transport models. However, in previous studies, new prediction algorithms have only been tested at stations or in a small region; a large-scale air quality forecasting model remains lacking to date. Huge dimensionality also means that redundant input data may lead to increased complexity and therefore the over-

fitting of machine learning models. Feature selection is a key topic in machine learning development, but it has not yet been explored in atmosphere-related applications. In this work, a regional feature selection-based machine learning (RFSML) system was developed, which is capable of predicting air quality in the short term with high accuracy at the national scale. Ensemble-Shapley additive global importance analysis is combined with the RFSML system to extract significant regional features and eliminate redundant variables at an affordable computational expense. The significance of the regional features is also explained physically. Compared with a standard machine learning system fed with relative features, the RFSML system driven by the selected key features results in superior interpretability, less training time, and more accurate predictions. This study also provides insights into the difference in interpretability among machine learning models (i.e., random forest, gradient boosting, and multi-layer perceptron models).

来源：GMD

发布时间：2022-10-24

数据类型：期刊

https://gmd.copernicus.org/articles/15/7791/2022/