Загрязнение поверхностных вод в России: ежемесячные данные о высоком и экстремально высоком загрязнении водных объектов за 2008—2021 гг.

Лицензия: CC BY-SA
Доступ: Доступен всем пользователям
Дата обновления: 23.09.2021
Теги:
экология, охрана окружающей среды, Росгидромет, загрязнение поверхностных вод

Описание

Набор данных включает в себя ежемесячные сведения о загрязнении поверхностных водных объектов на территории России за период с 2008 по 2021 гг. Экстремально высокие загрязнения доступны с детализацией до муниципальных образований; высокие загрязнение – с детализацией до субъекта РФ. Эти сведения публикуются в ежемесячных отчетах Росгидромета на официальном сайте в формате docx-файлов. В ходе подготовки этого датасета, данные были собраны из этих файлов и преобразованы в машиночитаемые таблицы в формате панельных данных, а также обогащены дополнительными атрибутами.

Единица наблюдения в датасете — место, в котором зафиксирован случай или случаи высокого или экстремально высокого загрязнения водного объекта за отчетный период (месяц). Данные для каждого наблюдения приведены по следующим атрибутам: регион, населенный пункт, в котором или вблизи которого произошел случай, наименование загрязняющего вещества или другого индикатора загрязнения, концентрация загрязняющего вещества или значение иного индикатора, класс опасности загрязняющего вещества и др.

Набор состоит из двух основных таблиц и одной вспомогательной:

- extreme_pollution.csv (основная); 7671 наблюдений по 13 атрибутам за 01.01.2008 – 31.08.2021;
- high_pollution.csv (основная 10198 наблюдений по 11 атрибутам за 01.01.2008 – 31.08.2021;
- pdk.csv (вспомогательная): 212 наблюдений по 10 атрибутам.

Набор доступен для работы в формате CSV (кодировка: «UTF-8», разделитель: «;»).

Характеристики датасета

Использование

Лицензия: CC BY-SA
Доступ: Доступен всем пользователям

Предоставление данных

Формат:
CSV

Характеристики набора

Временной период: 2008—2021 гг.
Количество наблюдений: 18081

Обновления

Размещено: 23.09.2021
Обновлено: 23.09.2021
Частота обновления: не указано
Дата следующего обновления: не указано

Происхождение

Источники данных: Официальный сайт Росгидромета
Ответственные за набор данных: ,

Структура набора данных

Экстремально высокое загрязнение поверхностных вод (extreme_pollution)

Атрибут	Описание	Количество пропусков (NaN) (шт/%)	Единица измерения	ОКЕИ	Формат
id	ID наблюдения в рамках набора данных	0			integer
period	Последний день отчетного периода (месяца)	0			date
subject	Наименование субъекта Российской Федерации, в котором зафиксирован случай загрязнения	0			cat
municipality	Наименование муниципального образования Российской Федерации, в котором зафиксирован случай загрязнения	347			cat
settlement	Наименование населенного пункта, в котором или вблизи которого произошло загрязнение	161			cat
oktmo	Значение кода ОКТМО для населенного пункта. Код ОКТМО населенных пунктов состоит из четырех частей и 11 знаков: 2 знака — «код региона»; 3 знака — «код района/города МО»; 3 знака — «код поселения МО»; 3 знака — «код населенного пункта МО»	347			numeric
water_body	Наименование водного объекта, в котором зафиксировано загрязнение Например, «вдхр. Аргазинское», «оз. Большое Островное», «р. Колос-Йоки»	0			cat
indicator	Загрязняющее вещество, по которому превышен ПДК, или иной показатель, значение которого не соответствует нормативу, что характеризует ситуацию как экстремально высокое загрязнение. Например, «Аммоний-ион», «Взвешенные вещества», «БПК5», «Водородный показатель (pH)»	0			cat

hazard_class	Класс опасности загрязняющего вещества: 1-й класс — чрезвычайно опасные; 2-й класс — высоко опасные; 3-й класс — опасные; 4-й класс — умеренно опасные; 4-й Э — «экологический» подкласс 4 класса. Для иных индикаторов, которые не являются загрязняющими веществами (например, запах, БПК5 или pH), класс опасности не указан, так как не имеет смысла	2538			cat
cnt_cases	Количество случаев загрязнения, зафиксированных в данном месте в данном месяце	0			integer
value_min	Минимальная измеренная концентрация загрязняющего вещества или минимальное зафиксированное значение иного индикатора из всех случаев, зафиксированных в данном месте в данное время	1			numeric
value_max	Максимальная измеренная концентрация загрязняющего вещества или максимальное зафиксированное значение иного индикатора из всех случаев, зафиксированных в данном месте в данное время	1			numeric
unit	Единица измерения значения, указанного в полях value_min, value_max: ПДК — число раз, в которое зафиксированная концентрация загрязняющего вещества превышает его ПДК, действующий в данном месяце данного года. Иначе говоря, это частное от деления измеренной концентрации загрязняющего вещества на его ПДК; мг/л — миллиграммы на литр; pH — единицы pH; балл — условная характеристика интенсивности запаха	0			cat

Справочная информация по ПДК загрязняющих веществ (pdk)

Атрибут	Описание	Количество пропусков (NaN) (шт/%)	Единица измерения	ОКЕИ	Формат
date_end	Дата завершения действия нормативно-правового акта, указанного в поле npa	0			date

hazard_class	Класс опасности загрязняющего вещества: 1-й класс — чрезвычайно опасные; 2-й класс — высоко опасные; 3-й класс — опасные; 4-й класс — умеренно опасные; 4-й Э — «экологический» подкласс 4 класса	1			cat
value	Предельно допустимая концентрация (ПДК)	0			numeric
formula	Формула вещества	26			text
n_cas	Регистрационный номер CAS	97			text
npa	Сокращенное название нормативно-правового акта (НПА), из которого взята информация	0			cat
date_start	Дата начала действия нормативно-правового акта, указанного в поле npa	0			date
ingredient	Название загрязняющего вещества (нормализованное)	0			text
ingredient_npa	Название загрязняющего вещества (исходное из нормативно-правового акта)	0			text

lpv	Лимитирующий показатель вредности (ЛПВ): токс — токсикологический (прямое токсическое действие веществ на водные биологические ресурсы); сан-токс — санитарно-токсикологический (действие вещества на водные биологические ресурсы и санитарные показатели водных объектов рыбохозяйственного значения); общ — общесанитарный; орг — органолептический (образование пленок и пены на поверхности воды водных объектов рыбохозяйственного значения, появление посторонних привкусов и запахов в воде водных объектов рыбохозяйственного значения, выпадение осадка, появление опалесценции, мутности и взвешенных веществ, изменение цвета воды водных объектов). При этом указывается расшифровка характера изменения органолептических свойств воды водных объектов рыбохозяйственного значения: рыб-хоз — рыбохозяйственный (изменение товарных качеств видов водных биологических ресурсов, отнесенных к объектам промышленного и прибрежного рыболовства: появление неприятных и посторонних привкусов и запахов); зап. — изменяет запах воды; мутн. — у	0			cat

Высокое загрязнение поверхностных вод (high_pollution)

Атрибут	Описание	Количество пропусков (NaN) (шт/%)	Единица измерения	ОКЕИ	Формат
id	ID наблюдения в рамках набора данных	0			integer
period	Последний день отчетного периода (месяца)	0			date
okato	Код ОКАТО субъекта Российской Федерации, в котором произошло загрязнение	0			integer
subject	Наименование субъекта Российской Федерации, в котором произошло загрязнение	0			cat

river_basin	Название крупной реки, в речном бассейне которой зафиксировано загрязнение, либо указание на то, что загрязнение произошло в малой реке, озере, водохранилище или морской акватории. Например, «р. Амур», «малые реки, озера, водохранилища» и др.	0			cat
indicator	Загрязняющее вещество, по которому превышен ПДК, или иной показатель, значение которого не соответствует нормативу, что характеризует ситуацию как высокое загрязнение. Например, «Аммоний-ион», «Взвешенные вещества», «БПК5», «Водородный показатель (pH)»	0			cat
cnt_cases	Количество случаев загрязнения, зафиксированных в данном месте в данном месяце	0			integer
hazard_class	Класс опасности загрязняющего вещества: 1-й класс — чрезвычайно опасные; 2-й класс — высоко опасные; 3-й класс — опасные; 4-й класс — умеренно опасные; 4-й Э — «экологический» подкласс 4 класса. Для иных индикаторов, которые не являются загрязняющими веществами (например, запах, БПК5 или pH), класс опасности не указан, так как не имеет смысла	3303			cat
value_min	Минимальная измеренная концентрация загрязняющего вещества или минимальное зафиксированное значение иного индикатора из всех случаев, зафиксированных в данном месте в данное время	0			numeric
value_max	Максимальная измеренная концентрация загрязняющего вещества или максимальное зафиксированное значение иного индикатора из всех случаев, зафиксированных в данном месте в данное время	0			numeric

unit	Единица измерения значения, указанного в полях value_min, value_max: ПДК — число раз, в которое зафиксированная концентрация загрязняющего вещества превышает его ПДК, действующий в данный месяц данного года. Иначе говоря, это частное от деления измеренной концентрации загрязняющего вещества на его ПДК; мг/л — миллиграммы на литр; pH — единицы pH	0			cat

Внесенные в набор данных изменения по сравнению с источниками

Атрибуты из первоисточника по экстремально высокому загрязнению, связанные с территориальными единицами (названия субъектов РФ, муниципалитетов, населенных пунктов), были нормализованы согласно эталонному набору «Населенные пункты России: численность населения и географические координаты», опубликованному в каталоге ИНИД (https://data-in.ru/data-catalog/datasets/160/).
При этом:

- населенный пункт «г. Северский» (Свердловская область) нормализован как «г. Полевской»
  (г. Северский включен в состав г. Полевской Свердловской области в 1942 году);
- населенный пункт «пос. Молоканка» (Забайкальский Край) нормализован как «с. Абагайтуй»
  («Молоканка» — это территориальный пункт, входящий в состав села Абагайтуй).

В случае если сопоставить данные не удалось, наименование бралось из первоисточника с приведением формата к близкому к эталонному (например, с помощью удаления лишних пробелов). По некоторым наблюдениям не удалось восстановить код ОКТМО. Это связано с тем, что в первоисточнике есть населенные пункты, находящиеся в разных муниципальных образованиях, но на территории одного субъекта, и поскольку в наборе нет указания на муниципальное образование, возникает неоднозначность в идентификации принадлежности населенного пункта.

- В первоисточнике по экстремально высокому загрязнению встречались наблюдения, в которых зафиксировано несколько случаев загрязнения конкретным веществом в конкретном населенном пункте за отчетный период. В таблице ниже приведены три типовых ситуации.

Река, пункт	Регион	Ингредиент	Концентрация (ПДК)
р. Вязьма, г. Вязьма	Смоленская область	Кислород	0,2, 20 случаев
р. Омь, г. Омск	Омская область	Кислород, 4 случая	1,78; 2 по 0,93 и 0,46
р. Чапаевка, г. Чапаевск	Самарская область	Гексахлорциклогексан (ГХЦГ), 6 случаев	5–10

В первой ситуации строки разбивались на 20 наблюдений, значение 0,2 относилось и в value_min, и в value_max, при этом cnt_cases у каждого наблюдения заполнялось «1», т.е. один случай.

Во второй ситуации строки разбивались на 4 наблюдения; в первом из них value_min и value_max заполнялись значением 1,78, во втором и третьем — значением 0,93, в четвертом — 0,46; при этом cnt_cases у каждого наблюдения заполнялось «1», т.е. один случай.

Третья ситуация, поскольку в ней количество значений не совпадает с количеством случаев, оставлено одной строкой, при этом cnt_cases заполнялась значением «6», а value_min и value_max значениями «5» и «10» соответственно.

В наблюдениях, где зафиксирован единичный случай (таких подавляющее большинство), cnt_cases заполнялся значением «1», концентрация записывалась и в value_min, и в value_max. Как следствие, если cnt_cases равняется 1, то value_min = value_max.

- Набор данных дополнен таблицей, содержащей сведения о предельно допустимых концентрациях и классах опасности загрязняющих веществ согласно следующим нормативно-правовым актам:
  - Приказ Госкомрыболовства РФ от 28.04.1999 № 96 «О рыбохозяйственных нормативах» (Госкомрыболовство-99);
  - Приказ Росрыболовства от 18.01.2010 № 20 «Об утверждении нормативов качества воды водных объектов рыбохозяйственного значения, в том числе нормативов предельно допустимых концентраций вредных веществ в водах водных объектов рыбохозяйственного значения» (Росрыболовство-2010);
  - Приказ Минсельхоза России от 13.12.2016 № 552 «Об утверждении нормативов качества воды водных объектов рыбохозяйственного значения, в том числе нормативов предельно допустимых концентраций вредных веществ в водах водных объектов рыбохозяйственного значения» (Минсельхоз-2016);
  - «СанПиН 2.1.5.980-00. 2.1.5. Водоотведение населенных мест, санитарная охрана водных объектов. Гигиенические требования к охране поверхностных вод. Санитарные правила и нормы» (утв. Главным государственным санитарным врачом РФ 22.06.2000) (СанПиН-2000);
  - Постановление Главного государственного санитарного врача РФ от 30.04.2003 № 78 «О введении в действие ГН 2.1.5.1315-03» (вместе с «ГН 2.1.5.1315-03. Предельно допустимые концентрации (ПДК) химических веществ в воде водных объектов хозяйственно-питьевого и культурно-бытового водопользования. Гигиенические нормативы», утв. Главным государственным санитарным врачом РФ 27.04.2003) (СанПиН-2003);
  - Постановление Главного государственного санитарного врача РФ от 28.01.2021 № 2 «Об утверждении санитарных правил и норм СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» (вместе с «СанПиН 1.2.3685-21. Санитарные правила и нормы...») (СанПиН-2021).

- При извлечении информации из указанных выше нормативно-правовых актов было сделано следующее:
  - ПДК для нитрит-иона, нитрат-иона и аммоний-иона пересчитаны на чистый азот. Например, для нитрит-иона (NO2-) ПДК составляет 0,08 мг/л. С учетом массовой доли азота в нитрит-ионе, которая приблизительно равна 0,3, ПДК в пересчете на азот составит после округления 0,02 мг/л;
  - для фосфатов из трех доступных значений — для эвтрофных, мезотрофных и олиготрофных водоемов — выбрано значение ПДК для эвтрофных водоемов;
  - в качестве значения ПДК для ДДЭ (дихлордифенилдихлорэтилен) взято значение ПДК для позиции «Хлорорганические токсиканты, ДДТ и его метаболиты, ПХБ, альдрин, линдан и др.», так как ДДЭ является продуктом распада ДДТ;
  - для лигносульфонатов при анализе источника «Госкомрыболовство-99» в качестве значения ПДК было взято среднее значение ПДК для лигносульфонатов магния и лигносульфонатов натрия, так как отдельной позиции по лигносульфанатам в данном нормативно-правовом акте нет; класс опасности среди них был выбран более «жесткий» (3, а не 4).
- Наименования загрязняющих веществ и иных индикаторов загрязнения были нормализованы. Для нормализации названий использовались указанные выше нормативно-правовые акты, а также ежегодники «Качество поверхностных вод Российской Федерации», подготавливаемые Гидрохимическим институтом Росгидромета.
- Классы опасности загрязняющих веществ были проверены и уточнены по указанным выше нормативно-правовым актам. Если класс опасности был указан в нескольких одновременно действующих нормативно-правовых актах, то выбиралось значение из того, в котором была указана меньшая (более жесткая) предельно допустимая концентрация. Если ПДК были одинаковыми, то выбиралось меньшее (более жесткое) значение класса опасности. Класс опасности сульфатов (4) взят из ежегодника Гидрохимического института «Качество поверхностных вод Российской Федерации» за 2010 г., так как ни в одном нормативно-правовом акте он не указан.

Дизайн исследования

Читайте предложенные варианты для применения данного датасета на нашей wiki-странице

Росгидромет

Полное наименование источника данных	Ежемесячные отчеты Росгидромета об аварийном, экстремально высоком и высоком загрязнении окружающей среды, а также радиационной обстановке на территории России
Владелец источника данных	Федеральная служба по гидрометеорологии и мониторингу окружающей среды
Краткое описание источника данных	Отчеты размещены на официальном сайте Росгидромета. Отчет представляет собой один или несколько файлов в формате docx. В каждом отчете есть текстовая информация и таблицы. Среди таблиц есть одна с перечнем случаев экстремально высокого загрязнения поверхностных вод и одна с перечнем случаев высокого загрязнения поверхностных вод
Ссылка на источник данных	http://www.meteorf.ru/product/infomaterials/99/

Понятия, используемые в источнике	Высокое загрязнение (ВЗ) — состояние воды в водном объекте, соответствующее любому из следующих критериев: · максимальное разовое содержание для нормируемых веществ 1—2-й класса опасности в концентрациях, превышающих ПДК от 3 до 5 раз, для веществ 3—4-й класса опасности — от 10 до 50 раз (для нефтепродуктов, фенолов, соединений меди, железа и марганца — от 30 до 50 раз); · величина биохимического потребления кислорода (БПК5) от 10 до 40 мг/л; · снижение концентрации растворенного кислорода до значений от 3 до 2 мг/л; · покрытие пленкой (нефтяной, масляной или другого происхождения) от 1/4 до 1/3 поверхности водного объекта при его обозримой площади до 6 кв. км; · покрытие пленкой поверхности водного объекта на площади от 1 до 2 кв. км при его обозримой площади более 6 кв. км. Экстремально высокое загрязнение (ЭВЗ) — состояние воды в водном объекте, соответствующее любому из следующих критериев: · максимальное разовое содержание для нормируемых веществ 1—2-й класса опасности в концентрациях, превышающих ПДК в 5 и более раз, для веществ 3—4-й класса опасности — в 50 и более раз; · увеличение биохимического потребления кислорода (БКП5) свыше 40 мг/л; · снижение содержания растворенного кислорода до значения 2 мг/л и менее; · покрытие пленкой (нефтяной, масляной или другого происхождения) более 1/3 поверхности водного объекта при его обозримой площади до 6 кв. км; · покрытие пленкой поверхности водного объекта на площади 2 и более кв. км при его обозримой площади более 6 кв. км; · появление запаха вод интенсивностью более 4-х баллов и не свойственного воде ранее; · массовая гибель моллюсков, раков, лягушек, рыб, других водных организмов и водной растительности. Предельно допустимая концентрация (ПДК) — максимальная концентрация вредного вещества, при которой в водном объекте не возникает последствий, снижающих его рыбохозяйственную ценность или возможность использования для хозяйственно-питьевых целей. Можно выделить две разновидности ПДК: рыбохозяйственные и санитарно-гигиенические. Первые устанавливаются для водных объектов рыбохозяйственного значения, вторые — для водных объектов хозяйственно-питьевого и культурно-бытового назначения. Если для одного и того же вещества установлены разные ПДК, то при мониторинге окружающей среды используется наиболее «жесткое», то есть наименьшее, значение. Класс опасности — условная величина, характеризующая вредное воздействие вещества на здоровье человека или экологические системы. Есть четыре класса опасности: 1-й класс — чрезвычайно опасные вещества; 2-й класс — высоко опасные; 3-й класс — опасные; 4-й класс — умеренно опасные. Кроме того, при мониторинге окружающей среды в четвертом классе выделяют «экологический» подкласс 4-э. Он означает, что действие вещества проявляются в изменении экологических условий в водоеме. Лимитирующий признак вредности в воде — признак, характеризующийся наименьшей безвредной концентрацией вещества в воде

Методология и изменения методологии для источника данных	Порядок представления информации об аварийных выбросах вредных веществ установлен в Приказе Росгидромета от 31.10.2000 № 156 «О введении в действие Порядка подготовки и представления информации общего назначения о загрязнении окружающей природной среды»
Ссылки на методологию	https://docs.cntd.ru/document/901791258