



Объединение каталогов землетрясений: модифицированный метод ближайшего соседа для идентификации дублей

И.А.Воробьева, П.Н.Шебалин

План

- Постановка задачи и актуальность исследования
- Разработка метода ближайшего соседа для идентификации дублей при объединении каталогов землетрясений на примере афтершоковой последовательности землетрясения Тохоку.
- Практическое применение метода для создания каталога Восточного сектора Российской Арктики и каталога Южной Сибири.

Задача и актуальность исследования

- Каталоги землетрясений являются основой для множества фундаментальных исследований и прикладных задач
- Из-за различных конфигураций сети и методов обработки записей разные агентства могут регистрировать/пропускать разные события, что в итоге выражается в отсутствии отдельных событий в различных каталогах землетрясений. Объединение данных различных сейсмологических агентств позволяет получить более полный каталог для исследуемого региона.
 - **Объединение всех доступных каталогов землетрясений, предполагающее идентификацию дублированных событий.**
 - **Унификация оценок магнитуд в объединенном каталоге**

Каталоги землетрясений

Базовая информация: время и место возникновения события и его сила

	Date	TIME	Lat°	Lon°	Dep, km	Mag	
1	22.11.1969	23:09:38.86	57.67	163.51	25.6	7.5	MS_ISC
2	18.05.1971	22:44:41.70	63.93	145.96	1.5	7.0	MS_ISC
3	08.03.1991	11:36:31.27	60.828	167.075	16.5	6.6	MW_GCMT

- Каталоги бывают глобальные, региональные, локальные
- Часто сейсмичность одной и той же зоны может быть представлена в нескольких разных каталогах
- Каталоги могут содержать разный набор событий в силу различной конфигурации сетей и способов обработки сейсмических записей.
- Объединение данных различных сетей позволяет улучшить представительность каталога землетрясений.

Магнитуды

Сила землетрясения описывается магнитудой.

Существует много типов магнитуды, в российских каталогах используется энергетический класс

Оценки магнитуд могут существенно различаться в зависимости от типа и агентства

Идентификация дублей при объединении каталогов землетрясений.

Дубли – это записи в различных каталогах землетрясений, относящиеся к одному сейсмическому событию. При объединении каталогов дубли выглядят как близкие в пространстве и времени события.

Оконный метод:

- Все доступные данные объединяются в один каталог.
- В объединенном каталоге идентифицируются группы близких событий, после чего из каждой такой группы выбирается одно событие, которое включается в итоговый каталог, остальные события исключаются

Недостаток оконного метода:

- В силу естественного группирования землетрясений (афтершоковые последовательности, рои) в автоматическом режиме неизбежно будут исключаться естественно близкие события, которые не являются дублями.
- Альтернативой является **ручной анализ** каждой группы близких событий
 - Субъективное решение исследователя
 - Трудозатратность

Метод ближайшего соседа для идентификации афтершоков

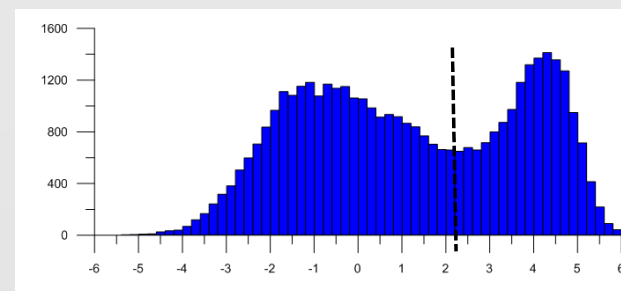
Zaliapin and Ben-Zion (2013, 2016)

Функция близости (Baiesi and Paczuski, 2004)

$$\eta_{ij} = \begin{cases} t_{ij}(r_{ij})^{d_f} 10^{-bm_i} & \text{for } t_{ij} > 0; \\ +\infty & \text{for } t_{ij} \leq 0 \end{cases}$$

Для каждого события ищется «ближайшее» предшествующее (порождающее) событие.

В итоге зависимыми считаем события с функцией близости меньше пороговой.



Распределение функции близости

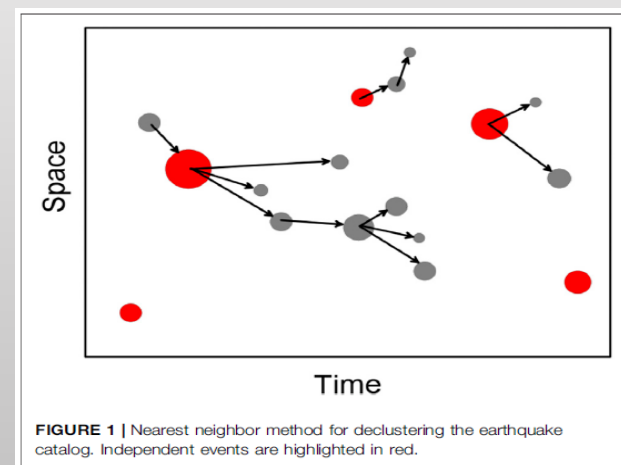


FIGURE 1 | Nearest neighbor method for declustering the earthquake catalog. Independent events are highlighted in red.

Модифицированный метод ближайшего соседа

Метод разработан по аналогии с современными способами идентификации афтершоков
Модификация метода ближайшего соседа основана на положении, что дубли формируют пары, в которых события необходимо принадлежат разным исходным каталогам.

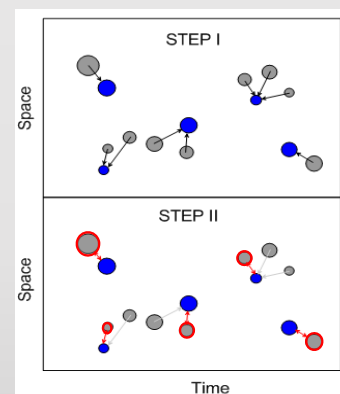
На входе имеется 2 каталога, главный каталог 1 и дополнительный каталог 2. Задача состоит в том, чтобы разделить каталог 2 на события, имеющие дубли в каталоге 1 и уникальные события.

Шаг 1. Для каждого события каталога 2 ищем ближайшего соседа из каталога 1 в соответствии с выбранной функцией близости. Таким образом, у каждого события из каталога 2 определяется единственное событие из каталога 1 для которого оно может являться дублем.

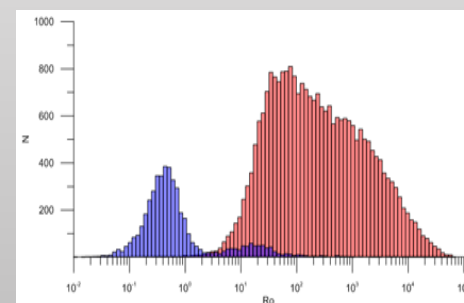
Шаг 2. Некоторые события из Каталога 1 могут оказаться ближайшими для нескольких событий Каталога 2, тогда в качестве потенциального дубля выбирается ближайшее из таких событий, остальные события объявляются не дублями, вне зависимости от значений метрики.

В итоге дублями считаем события второго каталога с функцией близости меньше пороговой. Остальные события Каталога 2 объявляются уникальными и добавляются к Каталогу 1.

Далее, можно последовательно добавлять любое число каталогов.



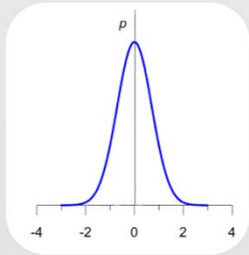
Синие – события из главного каталога 1
Серые- события из дополнительного каталога 2.
Потенциальные дубли выделены красным



Синяя гистограмма - распределение метрики для пар ближайших событий из сильно пересекающихся каталогов.

Функция близости для дублей: вероятностный подход

Мы предполагаем, что разница в определении землетрясения различными сетями является случайной величиной, имеющей нормальное распределение с нулевым средним по каждому из параметров



$$\begin{aligned}f(DT) &= \frac{1}{\sigma_T \sqrt{2\pi}} \exp\left(-\frac{DT^2}{2\sigma_T^2}\right), \\f(DX) &= \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left(-\frac{DX^2}{2\sigma_X^2}\right), \\f(DY) &= \frac{1}{\sigma_Y \sqrt{2\pi}} \exp\left(-\frac{DY^2}{2\sigma_Y^2}\right).\end{aligned}$$

DT, DX, DY – разница во времени, долготе и широте между различными определениями сейсмического события
 $\sigma_T, \sigma_X, \sigma_Y$ – соответствующие стандартные отклонения

Мы предполагаем, DT, DX, DY что являются независимыми случайными величинами, тогда вероятность дубля будет произведением вероятностей ошибок по каждому из параметров

$$f(DT, DX, DY) = \frac{1}{\sigma_T \sigma_X \sigma_Y (2\pi)^{3/2}} \cdot \exp\left(-\left(\frac{DT^2}{2\sigma_T^2} + \frac{DX^2}{2\sigma_X^2} + \frac{DY^2}{2\sigma_Y^2}\right)\right).$$

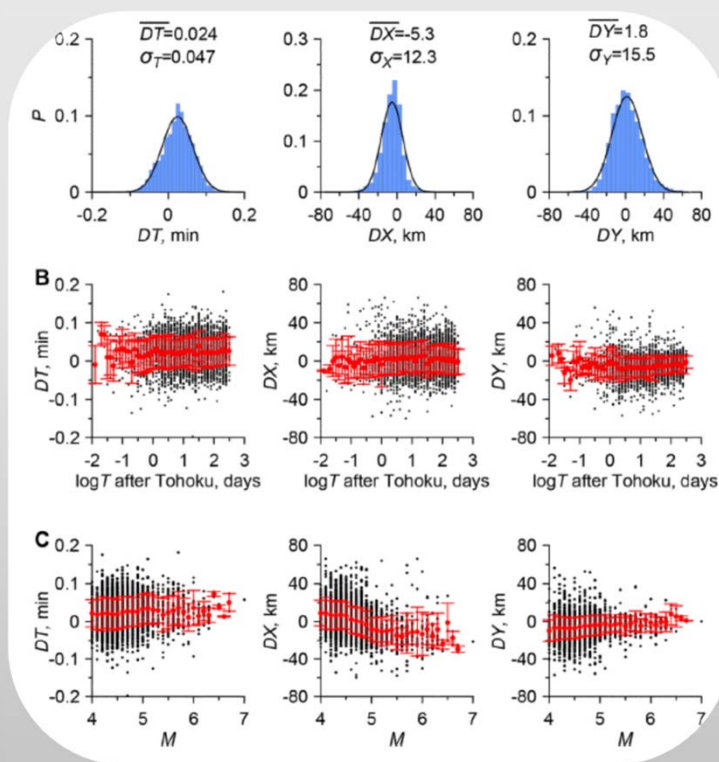
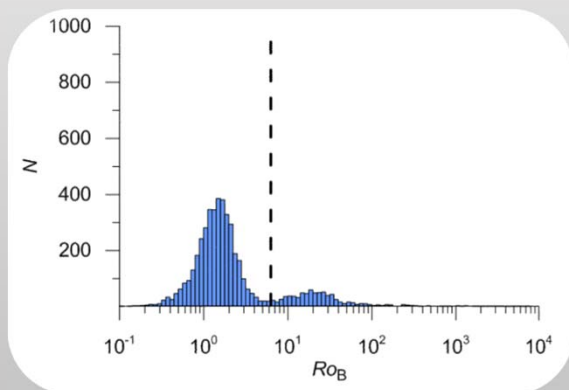
Таким образом мы естественно приходим к Эвклидовой метрике

$$Ro = \sqrt{\frac{DT^2}{\sigma_T^2} + \frac{DX^2}{\sigma_X^2} + \frac{DY^2}{\sigma_Y^2}}$$

Предварительная идентификация дублей и определение параметров метрики на примере афтершоков Тохоку из каталогов JMA и ANSS

Данные: 1 марта-31 декабря 2011, JMA $M \geq 3$, ANSS все события ($M \geq 4$)

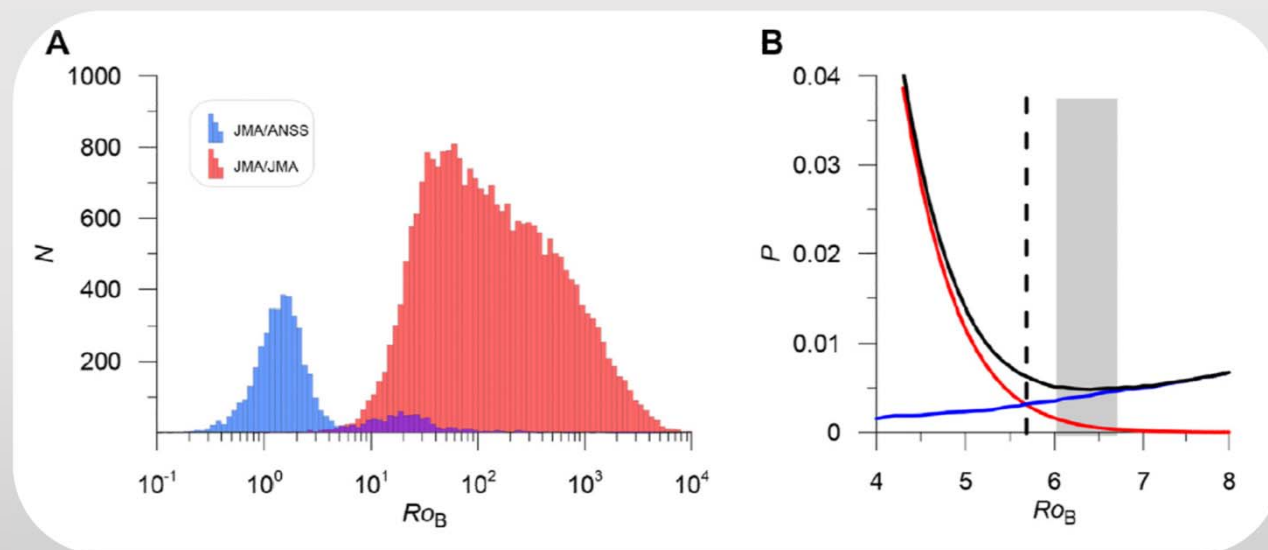
- Предварительный анализ дублей выполнен с параметрами распределения $\sigma_T = 0.05$ мин, $\sigma_X = \sigma_Y = 15$ км.
- Предварительное пороговое значение метрики выбирается в соответствии с минимумом распределения метрики



- Каждый из параметров DT , DX , DY следует нормальному распределению.
- Среднее мало по сравнению со стандартным отклонением для всех трех параметров DT , DX , DY .
- Дисперсия почти не зависит от магнитуды событий и времени

Идентификация дублей и определение порогового значения метрики на примере афтершоков Тохоку из каталогов JMA и ANSS

Данные: 1 марта-31 декабря 2011, JMA $M \geq 3$, ANSS все события ($M \geq 4$)



В. Оптимизация порога метрики

Красная линия – вероятность пропуска дубля в модели многомерного нормального распределения.

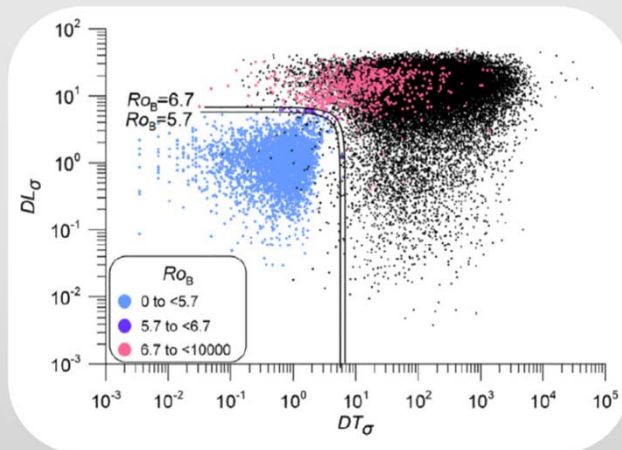
Синяя линия – вероятность ложного дубля

Черная линия – суммарная вероятность ошибок первого и второго рода

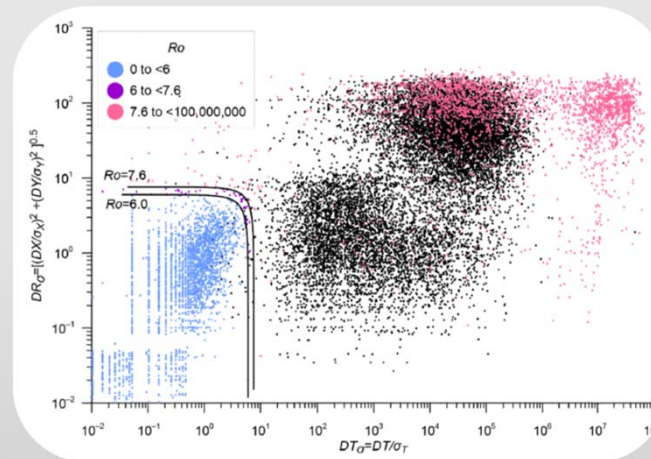
А. Сравнение распределения метрики для пар JMA/ANSS и JMA/JMA

Пространственно-временная структура дублей и афтершоков на примере каталогов афтершоков землетрясения Тохоку, Япония, и северо-востока Арктики

Тохоку ANSS/JMA



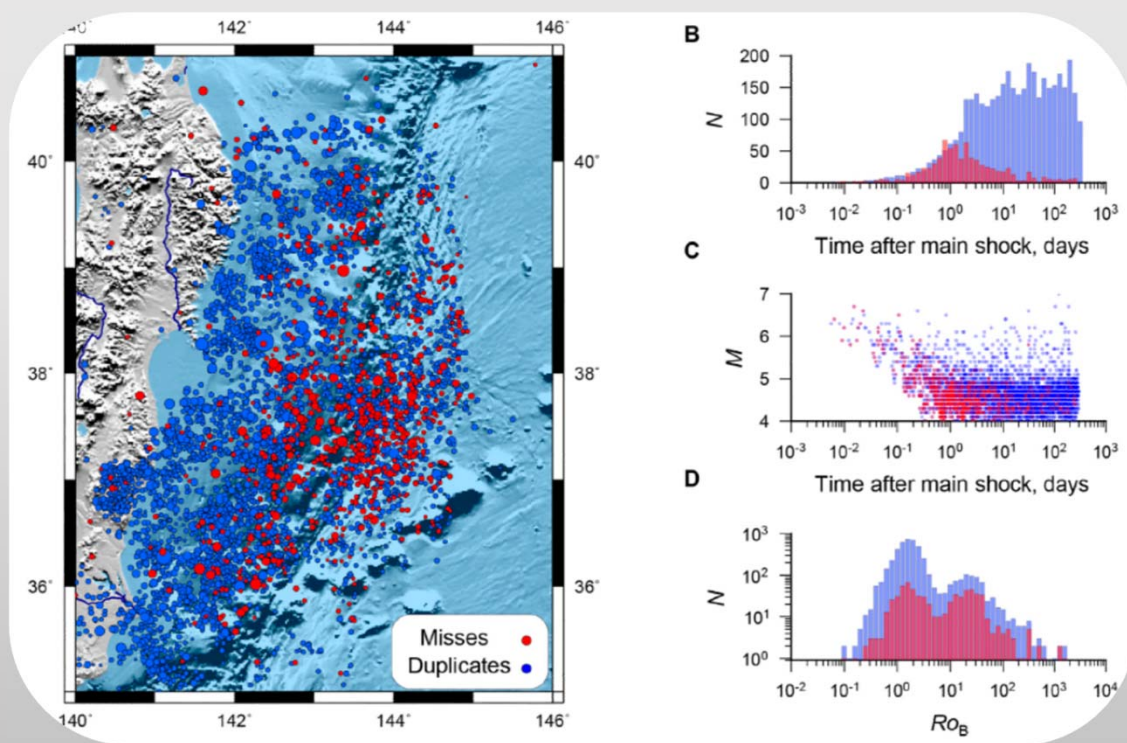
Арктика RUS/ISC



Распределение нормированных DT и DR , и линии уровня метрики Ro . Цветные точки – пары ,ближайших событий из разных каталогов, черные точки – расстояния между событиями в одном каталоге в метрике Ro . Значения метрики для землетрясений в одном каталоге существенно больше, чем для дублей из разных каталогов. Линии уровня метрики обеспечивают близкое к оптимальному разделение дублей и естественно группированных событий (нижний кластер черных точек).

Дубли и пропуски афтершоков Тохоку из каталогов JMA и ANSS

ANSS – 4709 событий. 759 (16%) событий отсутствуют в JMA. Надежность идентификации дублей 0.6%.



- A. Карта афтершоков из ANSS. Красным показаны события, пропущенные в JMA
- B. Распределение пропусков и дублей во времени. В первый день в JMA пропущена половина событий.
- C. Распределение магнитуды пропусков и дублей во времени
- D. Распределение метрики за весь период (синяя гистограмма) и в первый день (красная гистограмма)

Численные эксперименты

TABLE 1 | Summary of numerical experiments.

Experiment	Number of numerical parameters	Threshold value of metric	Number of duplicates	Number of events missed in JMA	Change in classification comparing with basic model, #/%	Error estimate (%)
Time period 01.03 2011–31.12.2011						
EX0 Basic model, metric Ro_B	3	5.7	3,950	759	—	.6
EX2 Symmetry test in basic model, metric Ro_B	3	5.7	3,950	759	0	.6
EX3 Model including magnitude, metric Ro_M	4	5.9	3,948	761	6/0.1	.6
EX5 Model including systematic errors and correlations, metric Ro_C	9	5.7	3,987	732	41/1	.6
First day after main shock						
EX1 Basic model, metric Ro_B	3	5.3	347	313	—	1.3
EX4 Model including magnitude, metric Ro_M	4	5.4	347	313	2/0.3%	1.4
EX6 Model including systematic errors and correlations, metric Ro_C	9	5.1	354	306	15/2.2%	1.9

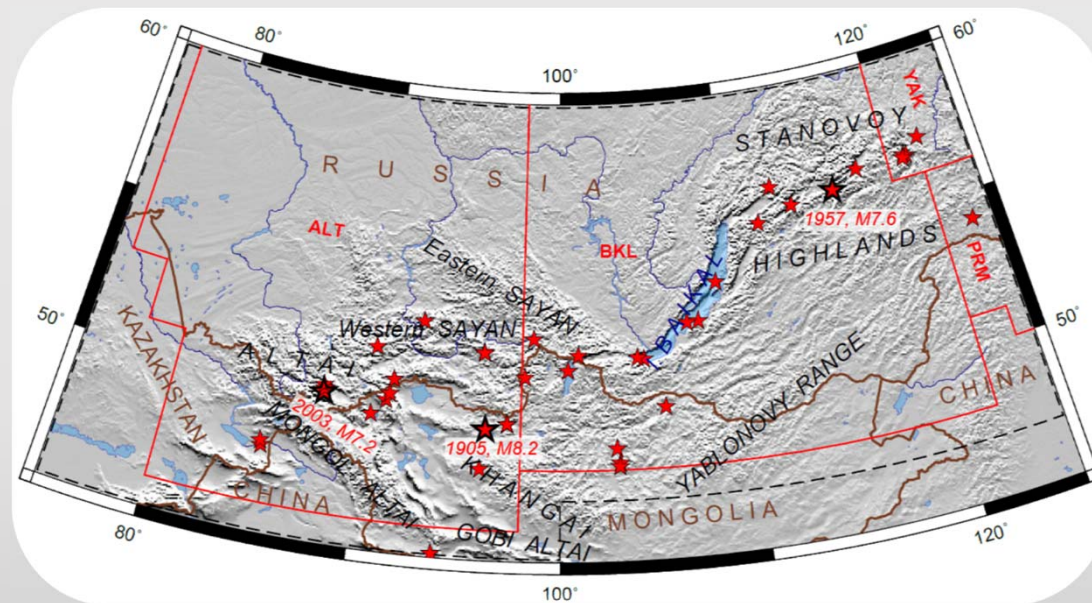
Выводы

- Метод идентификации дублей разработан по аналогии с современными методами идентификации афтершоков (Zaliapin and Ben-Zion, 2013; Zaliapin and Ben-Zion, 2016).
- Метод основан на вероятностной модели. Предполагается нормальное распределение разницы параметров события в разных источниках, что естественным образом приводит к Эвклидовой метрике в качестве функции близости.
- Метод позволяет оценить вероятность пропусков и ложных дублей.
- Процедура идентификации дублей автоматизирована и не требует дополнительного ручного анализа
- Метод предназначен для анализа современных инструментальных каталогов
- Базовая модель учитывающая разницу только во времени и координатах эпицентров достаточна для надежной идентификации дублей.

Применение методики для создания объединенных каталогов

- Сбор исходных каталогов
 - Приоритет источников данных (решение исследователя)
 - Порядок объединения данных с идентификацией дублей и оценками надежности
 - Анализ магнитуд
-
- Метод позволяет объединять любое число каталогов.
 - На каждом шаге объединяются 2 каталога, каждый из этих каталогов может быть результатом объединения данных на предыдущем шаге.
 - На каждом шаге объединения определен приоритет источников данных, так что событие с более высоким приоритетом автоматически попадает в итоговый каталог

Создание объединенного каталога землетрясений Южной Сибири



Регион исследования и зоны ответственности региональных отделений ГС РАН

Исходные данные: каталоги российских агентств и данные международного сейсмологического центра ISC

Российские данные

Землетрясения в СССР 1962-1991	4 региональных каталога
Землетрясения Северной Евразии 1992-2014	4 региональных каталога
Землетрясения России 2015-2019	4 региональных каталога
Региональный каталог Байкальского отделения ГС РАС 1994-2021	1 каталог
Итого:	13 каталогов, 72816 землетрясений $M \geq 2$ 16445 взрывов $M \geq 2$

Данные международного сейсмологического центра ISC 1962-2021

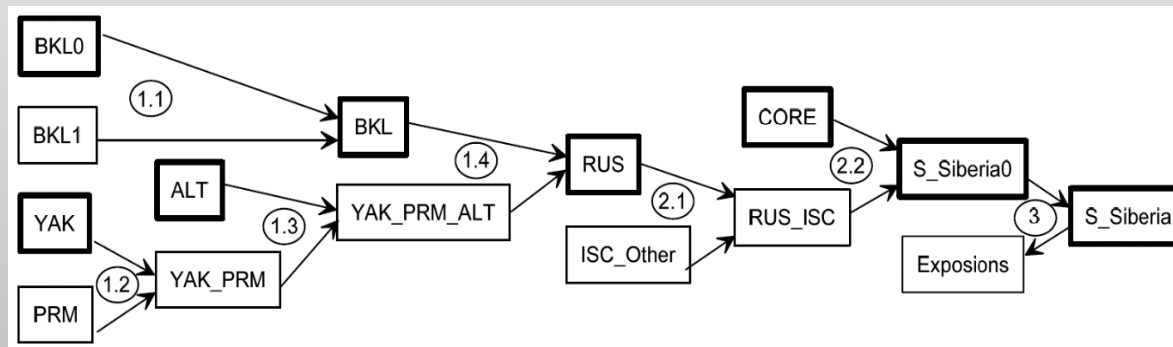
28 агентств	90849 землетрясений $M \geq 2$ 11432 взрыва $M \geq 2$
-------------	---

Приоритет источников данных и порядок сборки

1. События из глобального каталога ISC с магнитудами M_W^{GCMT} и/или mb^{ISC} CORE (2712 событий);
2. Землетрясения из российских каталогов. В зонах пересечения предпочтение отдается данным Байкала;
3. Остальные события из ISC, не включенные в CORE (без магнитуд M_W^{GCMT} и mb^{ISC}), (88137 событий). ISC_Other.

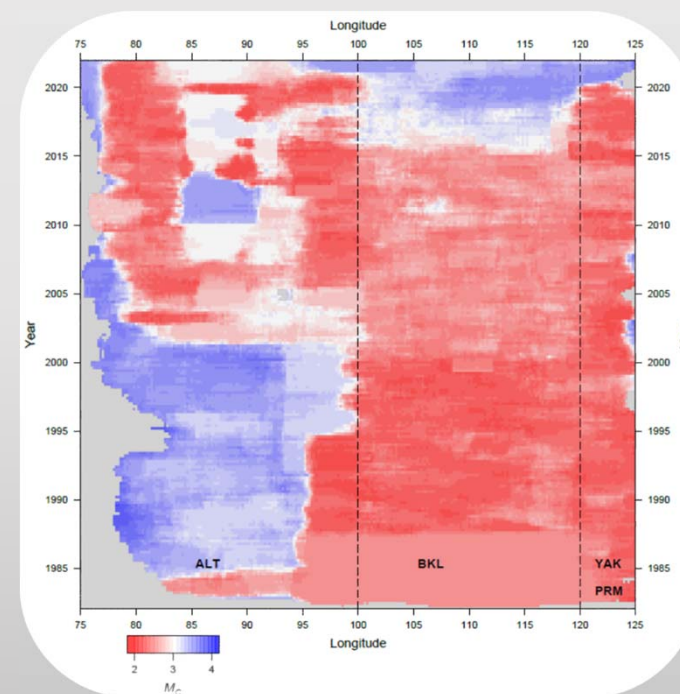
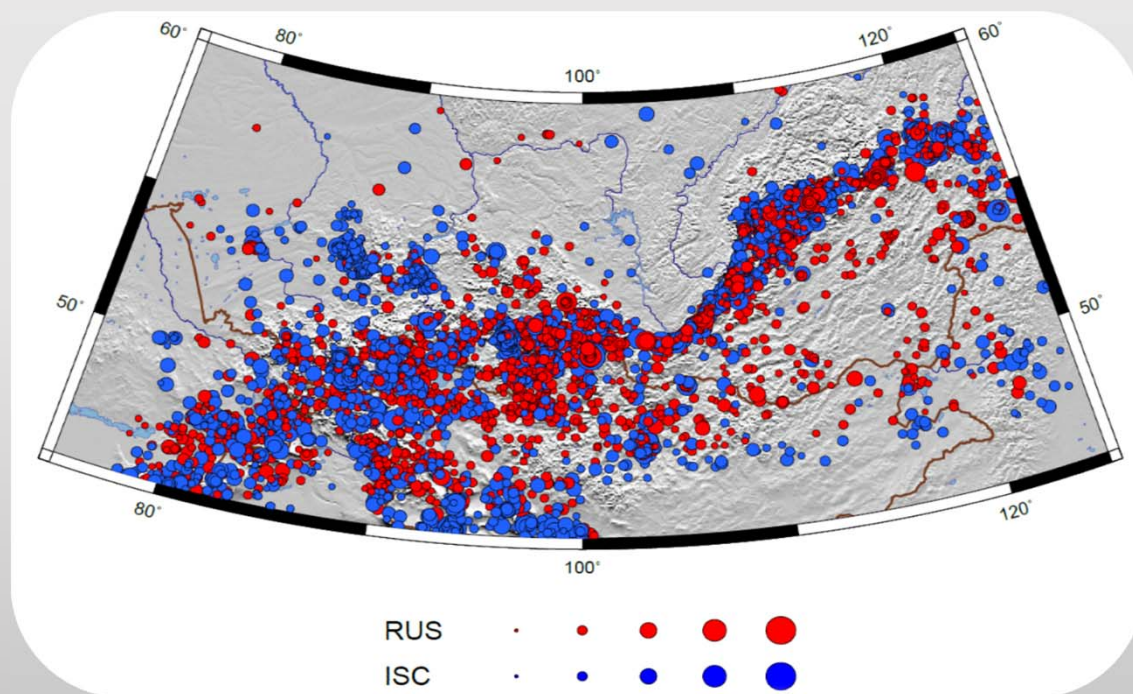
Итого объединяется 7 каталогов: 5 российских каталогов и два подмножества ISC.

Порядок сборки каталога



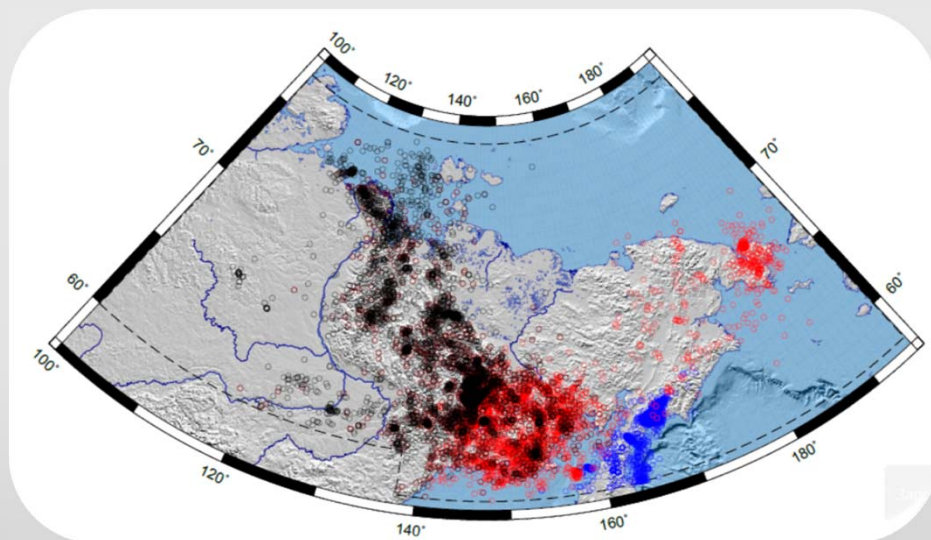
Карта эпицентров объединенного каталога Южной Сибири

Каталог содержит 112385 записей, 59701 из российских каталогов и 52684 записей из ISC.



Создание объединенного каталога землетрясений восточного сектора российской Арктики

Регион исследования



Эпицентры землетрясений из каталогов ГС РАН Якутии (черные), Северо-Востока (красные) и Камчатки (синие)

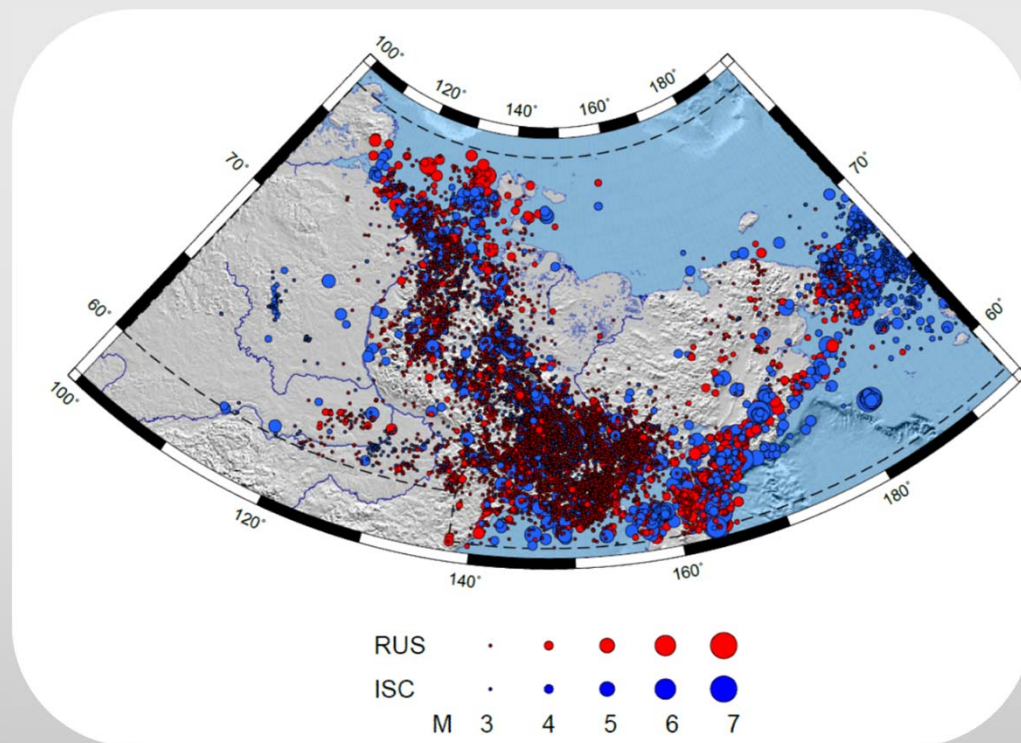
Каталоги

1. Региональный каталог Якутии из сборников «Землетрясения в СССР» 1962 – 1991, «Землетрясения Северной Евразии» 1992 – 2014 и «Землетрясения России» 2015 – 2019 (ГС РАН) (далее YAK).
2. Региональный каталог Северо-востока России из сборников «Землетрясения в СССР» 1968 – 1991, «Землетрясения Северной Евразии» 1992 – 2014 и «Землетрясения России» 2015 – 2019 (ГС РАН) (далее NER).
3. Региональный каталог землетрясений Камчатки Камчатского филиала ГС РАН (далее KAM).
4. Каталог ISC, являющийся композитным каталогом, в котором собраны данные множества мировых агентств, в том числе российских. На территории исследования представлены данные 24 агентств

Приоритет источников данных

1. Данные глобальных сетей из каталога ISC
2. Данные Российских региональных сетей
3. Остальные данные из каталога ISC

Объединенный каталог землетрясений восточной Арктики



- На основе обобщения и интеграции данных различных сетей для восточного сектора Арктической зоны Российской Федерации составлен объединенный каталог землетрясений.
- Каталог содержит информацию о 23 254 сейсмических событиях за период 1962 – 2020 гг.
- 7 781 событие из ISC
- 15 473 события из Российских каталогов.

Магнитуды в объединенном каталоге землетрясений восточной Арктики

Объединенный каталог восточного сектора АЗРФ содержит 23 254 события, которые имеют разные типы магнитуды, определенные различными агентствами. Необходимо их унифицировать

Каталог ISC

Для каждого события приводятся все известные оценки магнитуды от множества агентств, событие может иметь несколько десятков различных магнитуд.

Каталоги ГС РАН

Большинство событий имеет энергетический класс. Число землетрясений с магнитудами невелико.

В каталогах Якутии и Северо-Востока приводятся классы Раутиан $kr = 1.8 \lg A_{\text{peak}} + f(r)$

В каталоге Камчатки приводятся классы Смирнова $ks = 2 \lg A_{\text{peak}} + f(r)$

Необходимо выбрать приоритет магнитуды (тип и агентство)

Необходимо установить корреляционные соотношения различных типов магнитуды.

Приоритет магнитуды

На сегодняшний день физической шкалой магнитуд является магнитуда по сейсмическому моменту M_W , которая является предпочтительной для оценки размера землетрясения в широком диапазоне. Поэтому все магнитуды будут преобразованы в квази M_W

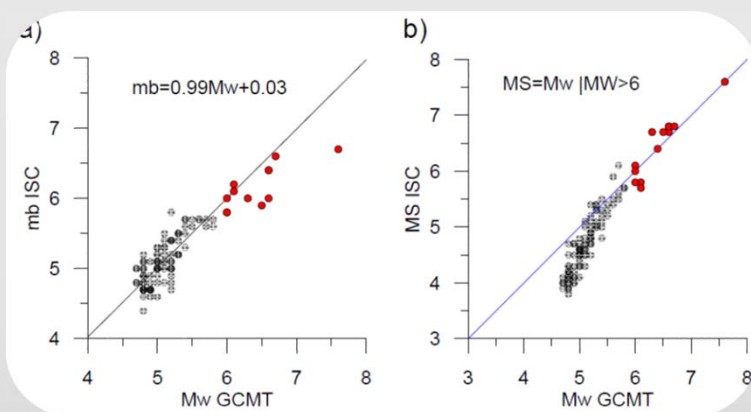
Рутинное определение M_W по данным глобальных сетей началось с 1976г. (каталог Global Centroid Moment Tensor, Harvard USA) для землетрясений с $M > 5.0$. При переходе к региональным оценкам повсеместно наблюдаются расхождения оценок M_W при $M < 5.0$, поэтому мы используем только глобальные оценки M_W GCMT.

В среднем диапазоне магнитуды мы отдаем предпочтение магнитуде mb_ISC , которые ISC использует в своей практике для получения оценок “квази M_W ” в диапазоне $M < 5.0$.

Для слабых землетрясений предпочтение отдается энергетическому классу, который известен для подавляющего большинства событий региона.

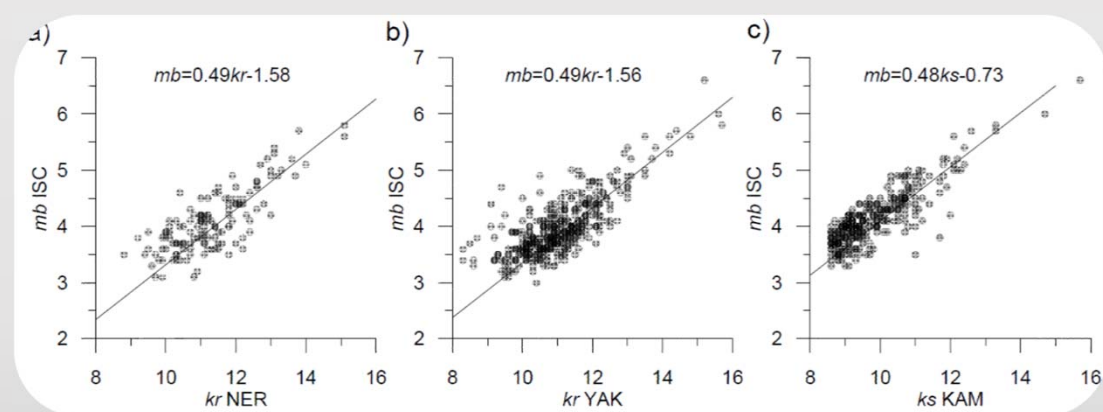
1. **M_W GCMT или MS ISC для сильных землетрясений до 1976 г. (107 событий);**
2. **mb ISC (1287 событий);**
3. **Магнитуда по энергетическому классу (20283 события);**
4. **Другие магнитуды (1575 событий). Всего использовано 27 типов различных магнитуд.**
Предпочтение отдается магнитудам, для которых можно построить соотношение с mb ISC
Корреляционные соотношения не определяются для 349 событий (1.5%)

Регрессионные соотношения магнитуд Арктика



Регрессионные соотношения магнитуд
GCMT и ISC:

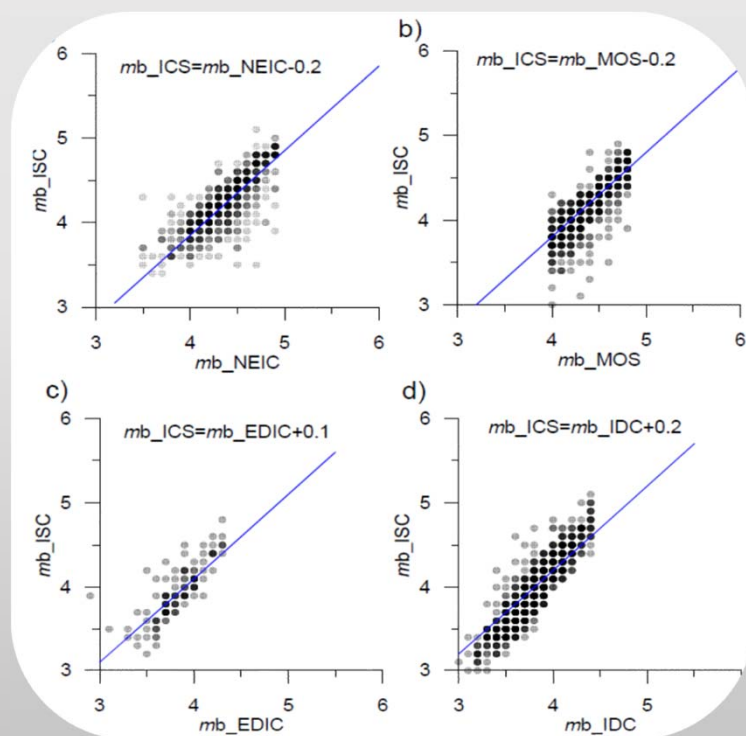
- a).** M_w GCMT и mb ISC для $M_w < 6.0$;
- b).** M_w GCMT и MS ISC для $M_w \geq 6.0$.



Регрессионные соотношения энергетических классов kr , ks и
магнитуды mb ISC.

- a).** Северо-Восток, kr ; **b).** Якутия, kr ; **c).** Камчатка, ks .

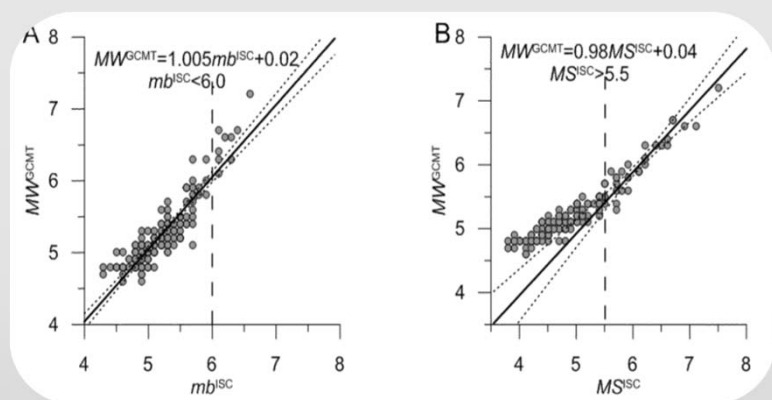
Регрессионные соотношения магнитуд m_b разных агентств в каталоге Арктики



Регрессионные соотношения магнитуды m_b
различных агентств и магнитуды m_b ISC.

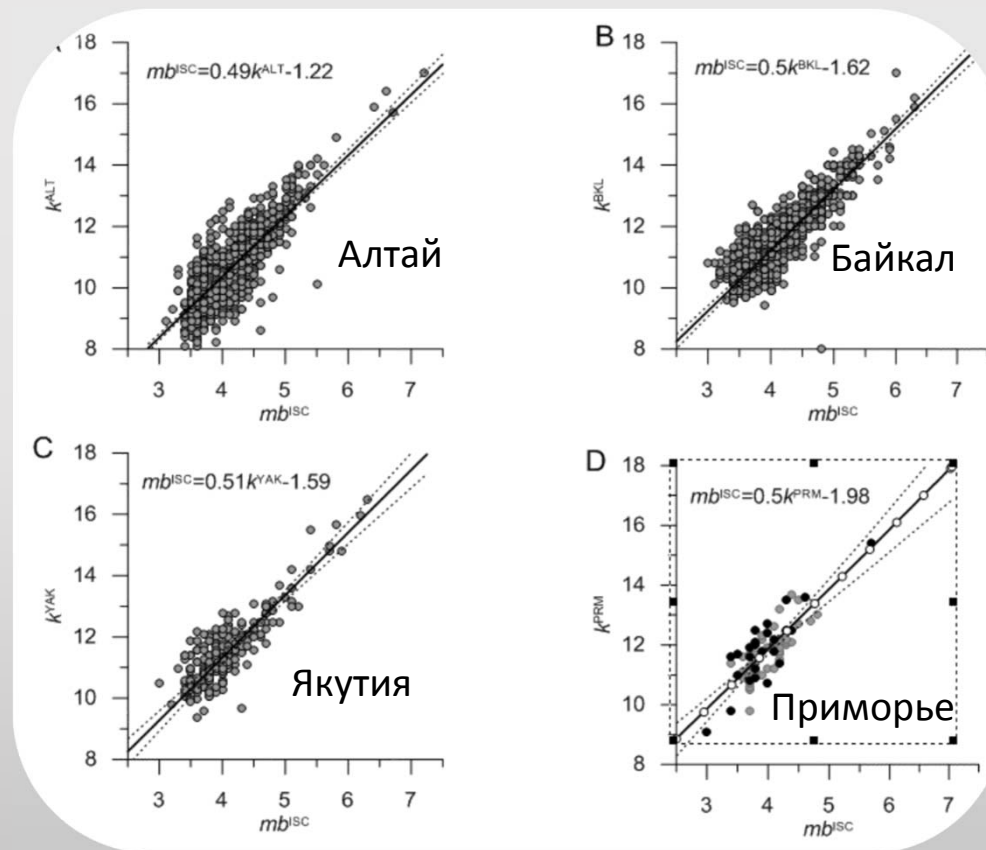
- a) National Earthquake Information Center, USA (NEIC);
- b) ГС РАН, Россия (MOS)ж
- c) Experimental International Data Center USA (EDIC);
- d) International Data Centre, Austria IDC.

Регрессионные соотношения магнитуд в каталоге Южной Сибири



Регрессионные соотношения магнитуд
GCMT и ISC:

- a). Mw GCMT и mb ISC для $Mw < 6.0$;
- b). Mw GCMT и MS ISC для $Mw \geq 6.0$.

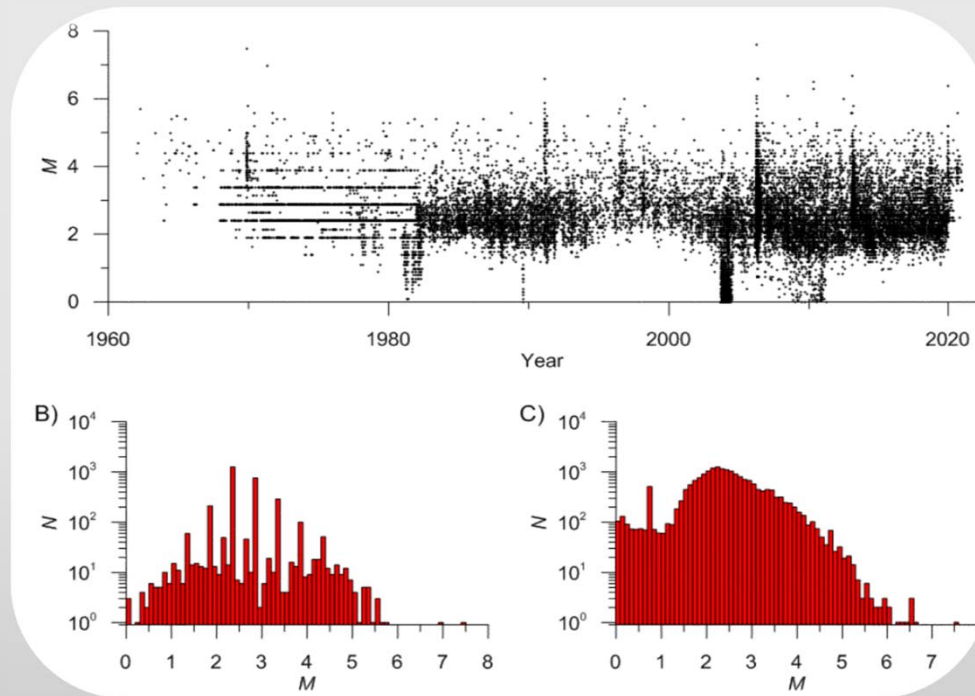


Регрессионные соотношения энергетических классов
и магнитуды mb ISC.

Заключение

- На основе интеграции данных различных агентств составлены каталоги землетрясений для восточного сектора Российской Арктики и Южной Сибири.
- Автоматическая идентификация дублей выполнена по авторской методике «Модифицированный метод ближайшего соседа» . Надежность идентификации дублей 1%
- Предложена унифицированная магнитудная шкала. Все оценки приведены к магнитуде по сейсмическому моменту на основе регрессионного анализа различных типов магнитуды.

Спасибо за внимание



A). Распределение магнитуды событий во времени;
B), C) Графики повторяемости до 1982 г. (**B**) и после 1982 г. (**C**).