

УДК 543.442.2

**РЕГУЛЯРИЗАЦИЯ МЕТОДОВ
БЕССТАНДАРТНОГО РЕНТГЕНОФАЗОВОГО АНАЛИЗА**

© 2011 И.С. Якимов*, П.С. Дубинин, А.Н. Залого, О.Е. Пиксина, Я.И. Якимов

*Сибирский федеральный университет, Красноярск**Статья поступила 22 марта 2010 г.**С доработки — 20 июня 2010 г.*

Сообщается о регуляризации "по Тихонову" метода ссылочных интенсивностей (RIR), интегрированного в составе информационно-поисковой системы рентгенофазового анализа с методом кластерной фазовой идентификации. Возможности методов обсуждаются на примере рентгенофазового анализа тестовых смесей из Round Robin on QPA, проведенного комиссией CPD международного союза кристаллографов.

Ключевые слова: качественный рентгенофазовый анализ, фазовая идентификация, нестандартный количественный рентгенофазовый анализ, метод ссылочных интенсивностей, метод RIR, регуляризация.

ВВЕДЕНИЕ

В последнее десятилетие был проведен ряд международных межлабораторных конкурсов (Round Robin) по рентгенофазовому анализу (РФА). Например, в конкурсе Search—Match Round Robin-2002 [1] требовалось определить качественный фазовый состав четырех образцов по выставленным в Интернет дифрактограммам. Дифрактограммы были запрошены 248 участниками, в финал вышло всего 30 участников, использовавших 10 различных компьютерных программ фазовой идентификации (ФИ), а среднее количество правильно идентифицированных ими фаз составило ~70 %. Проблемой ФИ материалов *сложного фазового состава* является плохое соответствие эталонных спектров фаз базы данных (БД) с экспериментальной дифрактограммой и отбор, вследствие этого множество спектрально подобных фаз-претендентов вместо однозначного результата [2]. В конкурсе Round Robin on QPA CPD IUCr [3] результаты КРФА восьми трехфазных смесей минералов характеризуются стандартным отклонением около 3 мас.%/фазу. В конкурсе Reynolds Cup competition [4] требовалось с помощью КРФА и любых других методов фазового анализа определить состав трех 12-фазных смесей глинистых и неглинистых минералов. Результаты количественного анализа характеризуются стандартным отклонением около 5 мас.%/фазу. Таким образом, объективные результаты Round Robin показали, что методы РФА еще далеки от совершенства и имеют серьезные проблемы при анализе многофазных материалов сложного фазового состава.

В данной работе представлен подход, основанный на интеграции методов ФИ и нестандартного КРФА в составе информационно-поисковой системы РФА (ИПС РФА) и применении регуляризации "по Тихонову" [5]. ИПС РФА первоначально развивалась как система фазовой идентификации [2], основанная на проблемно-ориентированном языке запросов [6, 7], специально разработанном для ФИ материалов сложного фазового состава, и средствах автоматического программирования запросов, встроенных в графический интерфейс пользователя (GUI ИПС). Критерием достоверности результатов ФИ в ИПС является модельный спектр дифрактограммы, представляющий линейную комбинацию идентифицированных эталонных спектров

* E-mail: I-S-Yakimov@yandex.ru

БД, например PDF2 ICDD, оптимизируемую по методу наименьших квадратов (МНК). В [8] описано использование регуляризации при генерации модельных спектров, включающих спектрально подобные эталоны БД и, следовательно, образующих плохо обусловленную систему уравнений МНК. В представленной работе, во-первых, описывается метод кластерной идентификации на базе языка запросов ИПС РФА, обеспечивающий выбор эталонного спектра истинной фазы из кластера подобных спектров фаз-претендентов и повышающий достоверность ФИ и устойчивость МНК. Во-вторых, предлагается новый способ регуляризации модельных спектров, обеспечивающий модернизацию классического метода ссылочных интенсивностей (Reference Intensity Ratio — RIR) [9] в новый вариант — мультирефлексный регуляризованный метод ссылочных интенсивностей (Regularized RIR — RRIR) [10]. Возможности методов обсуждаются на примерах рентгенофазового анализа тестовых смесей вышеупомянутых Round Robin.

РЕГУЛЯРИЗАЦИЯ РФА

Определим модель экспериментального спектра идентифицированного многофазного образца согласно [8] как линейную комбинацию:

$$I^{\text{mod}} = \sum S_j I_j \approx I_0, \quad j = 1, \dots, M, \quad (1)$$

где I_0 и I_j — векторы интенсивностей экспериментального спектра образца и j -го идентифицированного ссылочного спектра БД соответственно; S_j — коэффициенты приведения, которые можно определить по МНК в результате минимизации функционала:

$$\Phi(\bar{S}) = \sum_i^N W_i \cdot [I_{0i} - \sum_j^M S_j \cdot I_{ij}]^2, \quad (2)$$

где W_i — весовые коэффициенты; M — число ссылочных спектров; N — число линий экспериментального спектра.

В [10] нами показано, что для определения фазовых концентраций по методу ссылочных интенсивностей RIR вместо максимальных линий можно использовать коэффициенты приведения ссылочных спектров из (1) и (2). Тогда расчет фазовых концентраций по нормализованному варианту метода RIR производится по формуле

$$C_j = (S_j / K_j) / (\sum_l^M S_l / K_l), \quad (3)$$

а по варианту RIR с внутренним стандартом — по формуле

$$C_j = S_j / S_{\text{st}} \cdot K_{\text{st}} / K_j \cdot C_{\text{st}}, \quad (4)$$

где K_j и K_{st} — корундовые числа из БД для фазы j и фазы внутреннего стандарта соответственно; C_{st} — концентрация фазы добавляемого в образец внутреннего стандарта.

В распространенном случае, когда имеются данные анализа полного или частичного количественного элементного состава образца, а в образце отсутствуют твердые растворы неизвестного состава, концентрации фаз могут использоваться для составления системы уравнений элементного баланса. Произведя замену фазовых концентраций по выражениям (3) или (4), можно доопределить функционал МНК (2) данными элементного состава. Регуляризируя доопределенный функционал, получим функционал МНК регуляризованного метода ссылочных интенсивностей (RRIR). Для нормированного варианта RRIR этот функционал примет вид

$$\Phi(\bar{S}) = \sum_i^N W_i [I_{0i} - \sum_j^M I_{ij} S_j]^2 + \beta \sum_k^L w_k [\sum_j^m (P_{kj} - E_k) / K_i \cdot S_j]^2 + \alpha \sum_j^M S_j^2, \quad (5)$$

а для варианта RRIR с внутренним стандартом —

$$\Phi(\bar{S}) = \sum_i^N W_i [I_{0i} - \sum_j^M I_{ij} S_j]^2 + \beta \sum_k^L w_k [\sum_j^M P_{kj} / K_j \cdot S_j - E_k / (C_{\text{st}} K_{\text{st}}) \cdot S_{\text{st}}]^2 + \alpha \sum_j^M S_j^2, \quad (6)$$

где E_k — концентрация элемента k в образце; P_{kj} — атомная доля элемента k в фазе j ; L — число элементов образца с известным количественным составом; w_k — весовые коэффициенты; α —

параметр регуляризации; β — весовой коэффициент, регулирующий вклад дифракционных и химических данных.

При отсутствии данных элементного анализа $\beta = 0$, в этом случае функционал (5) становится эквивалентным функционалу, предложенному в [8] для построения модельного спектра как критерия качественного фазового анализа. При решении (5) или (6) по МНК искомые коэффициенты приведения S_j зависят от α и β . В [10] нами предложен алгоритм определения квазиоптимальных значений параметров регуляризации α и β с использованием невязки $D(S, \alpha) = \|I_0 - I^{\text{mod}}(S, \alpha)\|$ и характеристик погрешности элементного анализа. Вычислив с его помощью коэффициенты приведения, рассчитываем концентрации фаз по (3) или (4). Заметим, что физический смысл введения параметра регуляризации α в функционалы (5) и (6) заключается в компенсации почти всегда существующих небольших систематических различий в интенсивности эталонных спектров фаз из БД относительно спектров фаз анализируемого материала.

КЛАСТЕРНАЯ ФАЗОВАЯ ИДЕНТИФИКАЦИЯ

Метод кластерной фазовой идентификации является развитием идей, заложенных в указатели Ханавальта и Финка PDF2 ICDD [11], и базируется на применении вышеупомянутого языка для программирования запросов на ФИ. Каждый запрос является интерпретируемой программой, состоящей из последовательности отборочных операторов ФИ. Формат запроса имеет вид:

IN = "Идентификатор входной выборки или операция над выборками";

"Последовательность отборочных операторов".

OUT = "Идентификатор выходной выборки".

Последовательность операторов может включать операторы отбора по химическому составу образца (элементам или фрагментам химических формул), по указанным дифракционным линиям и спектральным критериям соответствия эталонного и идентифицируемого спектров. Процесс ФИ рассматривается как последовательность идентификационных запросов, создающих выборки из БД эталонных спектров или из предыдущих выборок, в результате которой создается итоговая выборка, содержащая эталонные спектры идентифицированных фаз образца. Программы-запросы скрыты от пользователя и формируются автоматически с помощью системы автоматического программирования в процессе оперирования пользователем на графических образах дифрактограммы, эталонных спектров, периодической таблицы и других объектах GUI ИПС. Графический интерфейс пользователя ИПС и система автопрограммирования запросов позволяют интерактивно применить стратегии ФИ "Поиск по сильной неидентифицированной линии дифрактограммы" и "Поиск по неидентифицированному химическому элементу образца" [2] и их различные комбинации. Данные стратегии создают возможность интерактивного итерационного выполнения процедуры ФИ, обеспечивающей разбиение множества фаз-претендентов на последовательность кластеров таким образом, что на очередной итерации попадает в каждый кластер одна истинная фаза (либо ни одной, при неадекватном запросе). Соответствующий ссылочный эталонный спектр включается в модельный спектр дифрактограммы образца, оптимизируемый далее по (2), (5) или (6) с приближенным расчетом фазовых концентраций по (3) или (4). На промежуточных итерациях ФИ элементный состав в (5) и (6) не используется ($\beta = 0$), а разность между измеренным и рассчитанным из значений фазовых концентраций элементным составом может служить индикатором возможности применения стратегии ФИ "по неидентифицированному химическому элементу". После завершения ФИ выполняется окончательный КРФА по (5) и (3) или (6) и (4).

ПРИМЕР ПРИМЕНЕНИЯ МЕТОДА КЛАСТЕРНОЙ ФАЗОВОЙ ИДЕНТИФИКАЦИИ

Рассмотрим применение совместной процедуры кластерной ФИ и КРФА по методу RRIR на примере фазового анализа дифрактограммы 7-фазного образца искусственного боксита, использованной в вышеупомянутом Round Robin on QPA. В табл. 1 приведены статистические результаты КРФА дифрактограммы боксита, полученные участниками Round Robin. Следует

Т а б л и ц а 1

Статистические результаты КРФА дифрактограммы боксита, мас.%

Фаза	Формула	Содержание	Среднее	Отклонение	Стандартное отклонение
Гиббсит	Al(OH) ₃	54,9	37,57	-17,33	6,59
Бемит	AlO(OH)	14,93	20,34	5,41	2,19
Гетит	FeO(OH)	9,98	13,93	3,95	3,35
Гематит	Fe ₂ O ₃	10	14,58	4,58	4,67
Кварц	SiO ₂	5,16	7,01	1,85	1,22
Анатаз	TiO ₂	2	3	1	0,43
Каолинит	Al ₂ (Si ₂ O ₅)(OH) ₄	3,02	3,94	0,92	1,96
Среднее отклонение				4,97	2,92

отметить большие систематические отклонения КРФА, вызванные, очевидно, сильной текстурой гиббсита. Ниже приведен элементный состав образца боксита (который участникам Round Robin известен не был):

Элементный состав	Al	Fe	Si	Ti
мас.%	26,34	13,27	3,07	1,20

Кластерная идентификация дифрактограммы боксита выполнена по базе данных из 115 000 эталонных спектров неорганических фаз за 7 шагов, без использования данных об элементном составе образца. Для этого оказалось достаточным применение основной стратегии ФИ — поиск по сильным неидентифицированным линиям дифрактограммы. На рисунке представлены шаги кластерной идентификации, идентифицируемые линии указаны стрелками. В таблицах на каждом рисунке показан список фаз в формате БД PDF2 из верхней части кластера, отсортированного ИПС по спектральному критерию соответствия эталонных спектров с дифрактограммой. Истинная фаза выбрана из списка визуально, по лучшему сходству эталонного спектра с дифрактограммой, и подчеркнута. Для отбора первых шести кластеров на дифрактограмме отмечалась только одна, а для седьмого — две из оставшихся неидентифицированными линиями (каолинита). Идентифицированный фазовый состав представлен в табл. 2.

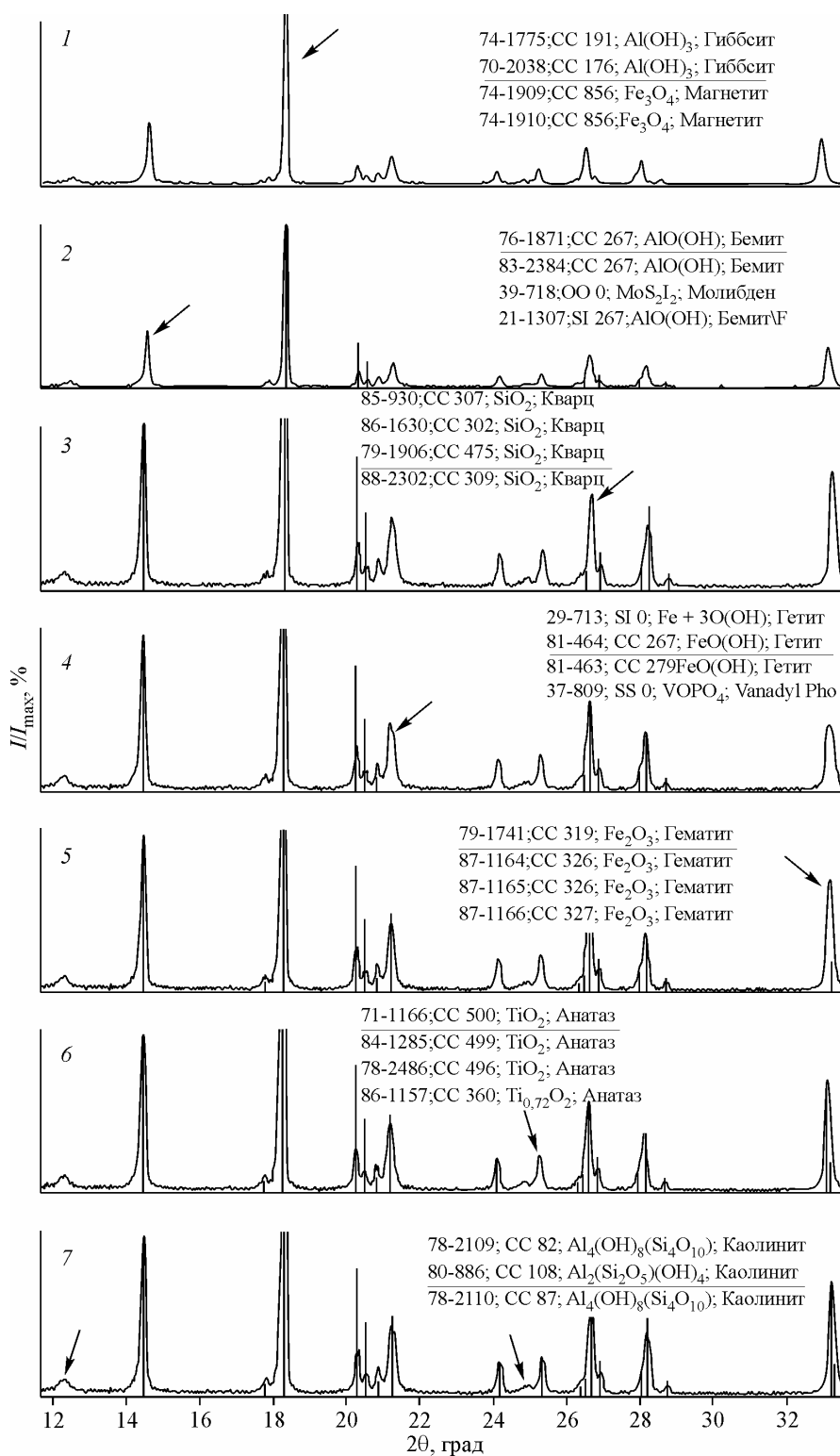
РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Каждая из минеральных фаз боксита, как и многих других неорганических фаз, представлена в БД несколькими эталонными спектрами, соответствующими различным структурным особенностям разных образцов данной фазы. Несмотря на большой объем БД в каждый кла-

Т а б л и ц а 2

Результаты КРФА боксита по методу кластерной ФИ—RRIR, мас.%

Фаза	Формула	Содержание	КРФА-1	Отклонение	КРФА-2	Отклонение
Гиббсит	Al(OH) ₃	54,9	55,6	0,7	54,2	-0,7
Бемит	AlO(OH)	14,93	14,4	-0,53	15	0,07
Кварц	SiO ₂	5,16	5,55	0,39	5,21	0,05
Гетит	FeO(OH)	9,98	9,32	-0,66	9,96	-0,02
Гематит	Fe ₂ O ₃	10	9,55	-0,45	10,1	0,1
Анатаз	TiO ₂	2	2,11	0,11	2,16	0,16
Каолинит	Al ₂ (Si ₂ O ₅)(OH) ₄	3,02	3,35	0,33	3,15	0,13
Среднее отклонение				0,46		0,19



Этапы кластерной идентификации дифрактограммы искусственного боксита

стер, как видно из рисунка, отбираются преимущественно эталонные спектры истинной фазы. Истинная фаза имеет, как правило, более высокую степень сходства с рентгенограммой и, следовательно, большее значение рейтингового вероятностного критерия соответствия по сравнению со спектрально подобными фазами-претендентами, случайно отобранными в кластер. Это

создает возможность выбора эталонного спектра истинной фазы из нескольких вариантов по наилучшему сходству с дифрактограммой с учетом тонких деталей соответствия. При этом и соответствующее корундовое число фазы, как правило, также оказывается наиболее адекватным. Как видно из состава кластеров, выбранный эталонный спектр не всегда имеет высшее значение рейтингового критерия. В результате модельный спектр дифрактограммы синтезируется наилучшим из возможных способов, а после его оптимизации по методу RRIR расчетные значения фазовых концентраций являются более точными. В табл. 2 представлены результаты КРФА по дифрактограмме боксита: в столбце КРФА-1 — для кластерной ФИ без данных об элементном составе, в столбце КРФА-2 — после уточнения модельного спектра по (5) с использованием данных об элементном составе образца.

Среднее отклонение расчетных фазовых концентраций от истинного фазового состава для КРФА-1 составляет $\sim 0,5$ мас.%, для КРФА-2 — всего $\sim 0,2$ мас.%. В последнем случае для уточнения по (5) использовался точный, вычисленный из табл. 1, количественный элементный состав металлов в образце, представленный выше.

Полученные результаты КРФА боксита по точности значительно превышают данные Round Robin on QPA, приведенные в табл. 1. Рассмотрим основные причины, приводящие к повышению точности КРФА описанным способом анализа. Во-первых, это отмеченный выше синтез модельного спектра из наиболее подходящих эталонных спектров с адекватными корундовыми числами, выбираемых в процессе кластерной ФИ. Во-вторых, основными сложностями дифрактограммы боксита для КРФА является сильная текстура гиббсита (см. рис., 2) и небольшая текстура бемита (см. рис., 3, линия справа от стрелки). При этом последняя определена по наилучшему в кластере эталонному спектру (№ 21-1307). После оптимизации модельного спектра по всем линиям дифрактограммы от этих фаз вклад текстуры в масштабные коэффициенты (5) усредняется, при этом линии эталонных спектров пропорционально понижаются. Для лучшего усреднения в (5) может применяться подходящая весовая схема. Наконец, в-третьих, отсутствие аморфной фазы (о чем можно судить, например, по минимальному уровню фона) позволило применить более точный, нормализованный вариант RRIR с расчетом фазовых концентраций по (3). В совокупности это позволило достичь весьма высокой точности анализа — со средним отклонением $\sim 0,5$ мас.%

Дополнительное использование точного количественного элементного содержания металлов в образце позволило улучшить точность КРФА до среднего отклонения $\sim 0,2$ мас.%. Следует отметить, что элементный состав металлов распределен по разным фазам и однозначно определяет только содержание анатаза. Повышение точности анализа остальных фаз обусловлено подгонкой концентраций групп фаз, содержащих тот или иной металл, к его содержанию при оптимизации модельного спектра по (5). При этом значение регуляризационного параметра β увеличивается до прекращения существенной сходимости расчетного содержания металлов к истинным (или измеренным) значениям. Очевидно, точный элементный состав образца — идеальный случай, и среднее отклонение $\sim 0,2$ мас.%. Следует рассматривать как предельную точность метода. Однако, как показано в [10], использование данных и не вполне точного элементного анализа заметно повышает точность КРФА. Дело в том, что при использовании элементного состава происходит доопределение функционалов (5) или (6) с соответствующим увеличением числа нормальных уравнений МНК, вследствие чего повышается сходимость и точность КРФА. Но взаимосвязь между величинами погрешности элементного и фазового анализа в этом направлении еще недостаточно изучена.

ВЫВОДЫ

Кластерный фазовый анализ 7-фазного искусственного боксита был выполнен за 7 шагов по базе данных из 115 000 эталонных спектров без использования данных об элементном составе образца. При этом достигнута точность нестандартного КРФА, характеризуемая средним отклонением от истинного состава $\sim 0,5$ мас.%/фазу. Метод кластерной фазовой идентификации существенно упрощает процедуру ФИ, совмещает ее с процедурой КРФА по методу RRIR и повышает достоверность идентификации и точность количественного анализа, которые могут

быть еще более повышены за счет использования данных количественного элементного анализа. Возможно, главным достоинством представленного подхода является следующее: у пользователя создается уверенность в том, что истинные фазы не были опущены из рассмотрения и действительно идентифицированы, а точность бесстандартного КРФА соответствует точности независимо выполненного элементного анализа.

Метод кластерной ФИ и некоторые варианты метода RRIR интегрированы в составе ИПС РФА, которая остается не только эффективным инструментом качественного фазового анализа сложных многофазных материалов, но и становится средством автоматического бесстандартного количественного анализа.

Представленная работа выполнена при поддержке гранта П2430 ФЦП "Научные и научно-педагогические кадры инновационной России" на 2009—2013 гг.

СПИСОК ЛИТЕРАТУРЫ

1. *Le Meins J.-M., Granswick L.M.D., Le Beil A.* // Powder Diffraction. – 2003. – **18**, N 2. – P. 106 – 113.
2. *Yakimov I.S.* // Inorgan. Mater. – 2008. – **44**, N 14. – P. 1531 – 1535.
3. *Scarlett N.V.Y., Madsen I.C., Cranswick L.M.D. et al.* // J. Appl. Crystallogr. – 2002. – **35**. – P. 383 – 400.
4. *Omotoso Oladipo, McCarty Douglas K., Hillier Stephen. et al.* // Clays and Clay Minerals. – 2006. – **54**, N 6. – P. 748 – 760.
5. *Тихонов А.Н., Арсенин В.Я.* Методы решения некорректных задач. – М.: Наука, 1979.
6. *Yakimov I.S.* // J. Nuclear instruments and methods in physics research. – 1989. – **A282**.
7. *Якимов И.С.* // Журн. Сиб. фед. ун-та. Сер. Хим. – 2009. – **2**, № 4. – С. 377 – 384.
8. *Burova E.M., Shchedrin B.M.* // Computational Mathematics and Modeling. – 2006. – **17**. N 1. – P. 45 – 55.
9. *Hubbard C.R., Evans E.H., Smith D.K.* // J. Appl. Crystallogr. – 1976. – **9**. – P. 169.
10. *Якимов И.С., Дубинин П.С., Пиксина О.Е.* // Журн. Сиб. фед. ун-та. Сер. хим. – 2009. – **2**, № 1. – С. 71 – 80.
11. PDF2-PDF4 Database. Version 2008. ICDD, US.